CHILD DEVELOPMENT

Vol. 29
March, 1958
No. 1

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

BETTYE MCDONALD CALDWELL
Washington University School of Medicine

ERCEL S. EPPRIGHT
Iowa State College

EUGENE S. GOLLIN
Queens College

HARRY LEVIN
Cornell University

PAUL H. MUSSEN
University of California

KATHERINE E. ROBERTS
Merrill-Palmer School

WATARU W. SUTOW
University of Texas

WILLIAM W. WATTENBERG
Wayne State University

Managing Editor

KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Notice to Contributors

Child Development is devoted to basic research and theory concerning the growth and development of the child.

Manuscripts should be submitted in duplicate, each copy bearing only the title as a means of identification. The title, author's name, and institutional affiliation (together with any footnotes which reveal the identification of the author) should be submitted on separate pages.

Contributors are asked to refer to the *Publication Manual* of the *American Psychological Association* (1957 Revision) in the preparation of manuscripts.

The following rules require special emphasis:

1. All copy, including indented matter, should be typed double-spaced on white standard paper. Lines should not exceed six inches in length.

2. Footnotes should be few in number. They should be indicated in the text by superscript arabic numerals, numbered consecutively throughout the manuscript. The footnotes themselves should be typed double-spaced on a separate page placed at the end of the manuscript.

3. Tables and figures are costly to print and should be used only when essential. The author will be asked to contribute to the cost of composition of tabular and graphic material, which is excessive in amount or unduly complicated in form. Each table should be typed on a separate sheet. Insert a guide line (e.g., Table 1 about here) at the appropriate place in the manuscript. Figures should be prepared with black India ink, suitable for photography. Legends typewritten on an ordinary typewriter are unacceptable.

4. References in the text should be indicated by numbers in parentheses that refer to the list of references at the end of the article. Do not use footnotes for this purpose. In the list of references, the arrangement should be alphabetical by author's surname, or for multiple authors by the name of the senior; for any one author, by order of date. The following standard form should be observed:

> Jones, H. E. Development in adolescence. New York: D. Appleton-Century, 1943.
>
> Shepard, Winifred O., & Schaeffer, M. The effect of concept knowledge on discrimination learning. Child Develpm., 1956, 27, 173-178.

The World List of abbreviations is used for journals.

The manuscript submitted should represent the final form in which the author wishes the paper to appear. All changes in proof (other than printer's errors) will be charged to the author.

All manuscripts should be addressed to William E. Martin, Editor, Child Development Publications, Purdue University, Lafayette, Indiana.

CONTRIBUTIONS OF RESEARCH WITH CHILDREN TO A THEORY OF MALADJUSTMENT[1]

JOEL R. DAVITZ
Teachers College, Columbia University

The major theories of psychological maladjustment, such as those of Freud and Horney, agree that childhood experience is of primary importance in the development of neurosis. It is remarkable, therefore, that almost all of these theories were originally formulated without reference to any but the most casual kind of research about children. In the main, the meanings of childhood experiences were interpreted by these theorists largely on the basis of inferences from observations of adult behavior. Only *after* the theories had been formulated did researchers try to find support for these chronologically "backward looking" inferences in observations of children. Unfortunately, these efforts have not been consistently successful. In fact, very little evidence supports most of the central hypotheses about childhood and maladjustment proposed by the major theories of neurosis. Consider, for example, the striking absence of reliable research evidence in support of the Freudian hypotheses about the Oedipus complex or about the effects of infantile frustrations on personality development. The absence of such support some 50 years after these hypotheses were published suggests that another approach to a theory of maladjustment might possibly result in a more substantial relationship between theory and research evidence. If we accept the premise that childhood experience is important in the development of maladjustment, rather than begin with observations of adults and infer back to childhood, perhaps it might be more useful to begin building a theory of maladjustment by starting with observations of children who are maladjusted.

The relationship between a child's early experiences with key figures in his environment and his subsequent personality development has been the focus of an impressive amount of research in child psychology. There are indeed large gaps of knowledge in this area, but the research available at the present time does permit some rather broad generalizations relevant

[1] Paper read at symposium on Contributions of Research with Children to Behavior Theory, Amer. Psychol. Ass., Chicago, September, 1956.

CHILD DEVELOPMENT, Vol. 29, No. 1 (March, 1958)

to a theory of maladjustment. Consider, for example, Symonds' studies (12) of parent-child relationships. On the basis of extensive research and a thorough review of the evidence, he concluded that if one or both parents strongly reject a child, he tends to become fearful, hostile, overaggressive, and may exhibit antisocial behaviors such as lying and stealing. In another series of researches, Redl and Wineman (8) reported that children who experienced severe rejection and abuse exhibit fear and anxiety, low frustration tolerance, irresponsibility, a tendency to panic when confronted with a novel situation, and greater difficulty in reacting to failure. Research also indicates that these negative reactions acquired in the home environment generalize to other interpersonal situations. Skeels and his co-workers (11) at Iowa, for example, in their study of environmental stimulation concluded that children reared in an orphanage without warm and affectionate parent-child relations tend to acquire attitudes of distrust and defensiveness which generalize in their relationships with others outside of the orphanage. And studies by Van Houten (13), Johnson (6), Hattwick and Stowell (4), and others have demonstrated that fearful responses acquired in the parent-child relationship tend to generalize to the school situation. From these and other related studies which have obtained similar results, we may conclude that in a punishing, rejecting interpersonal environment, the child tends to become fearful, and may exhibit hostility, overaggressiveness, distrust, irresponsibility, and other asocial or antisocial behaviors. Furthermore, the evidence suggests that these reactions tend to generalize and to influence the child's behavior in a wide range of interpersonal situations.

Up to this point we have been concerned with a straightforward stimulus-response course of learning. Punish, reject, abuse a child, and he will become fearful, suspicious, hostile. We are now faced with somewhat more complex problems. Although our research findings overwhelmingly support the common-sense generalization that threat engenders fear, not all children react fearfully to apparently equivalent punishing experiences. Some seem to blissfully live through years of threat and become happy, healthy persons. Others become maladjusted with seemingly minimal punishing experience. These contradictory observations present a crucial problem, and we do not have a satisfactory solution. But at least part of the solution may be found in studies of individual differences among infants in emotional sensitivity and responsiveness. The dimensions of temperamental differences in infancy, which are, presumably, genetically determined, have not yet been fully defined or measured; however, even at the present time there is a substantial body of research which supports the notion that such differences in sensitivity and responsiveness could well account for children's differential reactions to interpersonal threat. Shirley (10), for example, has described a wide range of differences in timorousness and irritability in early infancy that persist later in the child's life. Certainly it does not seem unreasonable to predict that the child who is

physically more sensitive, whose emotional reactions are cued off more easily and with greater intensity, will respond differently to stress in his interpersonal environment than will a child who is relatively insensitive to emotional stimulation.

A second problem involves the persistence of the child's emotional response long after the initial learning experience. Why doesn't the fear extinguish? This problem is what Mowrer (7) has termed the "neurotic paradox." A child whose parents are consistently rejecting and punishing learns to fear his parents and generalizes this fear to other parental figures. Sooner or later, however, one would expect this fear to decrease and in the normal course of events to extinguish. In the life of a psychologically healthy child, probably many fears are learned and then gradually extinguish. But in the maladjusted child, and adult, the fear of others persists for long periods of time without any apparent reduction or extinction. In fact, persistent, long-term fear might well be the chief defining characteristic of the maladjusted person.

The solution of this problem stems from the interpersonal nature of the child's learning experiences. The child lives and behaves in relationship to other people, and his behavior elicits some response from these other people. The infant's cry, for example, elicits some response from another person and this response may be either rewarding or punishing to the child. Sears (9) has developed this view of child development in his description of early socialization, emphasizing the interaction effects of parents and children behaving in relationship to each other.

Now consider the kinds of defensive behaviors found to be characteristic of children who have experienced severe punishment and rejection. Redl and Wineman (8) found hate, irresponsibility; Symonds (12) found hostility, overaggressivenesss, truancy, lying, stealing; other common characteristics that have been found are suspicion, severe withdrawal, various forms of delinquency, extreme attention-getting, and a long list of similar kinds of behaviors. What kinds of responses would one expect a child's persistent hostility, overaggressiveness, extreme attention-getting, severe withdrawal to elicit from other persons in his environment? In most cases, we could safely predict that the child's behavior will elicit some form of counterpunishment or rejection. Studies by Anderson (1), by Bishop (2), Brewer (3), and others support this prediction. Briefly, these researches have demonstrated that integrative behavior in one person tends to induce integrative behavior in others; domination tends to incite domination; aggression tends to be met with aggression. Anderson (1), in his review of the research with children in regard to this phenomenon, has called this "circular behavior." Jackson and Todd (5) in their discussion of maladjusted children have called this the "vicious circle."

The answer, therefore, to the question of why the maladjusted child's fear doesn't extinguish is simply that the fear continues to be reinforced. The sequence of events resulting in continued reinforcement of the child's

fear begins when the child learns to fear others as a consequence of rejection and punishment. Having learned to fear other persons, he develops defensive reactions that temporarily reduce his fear; but, in the long run, these defensive reactions elicit responses from other persons which are punishing and which reinforce his fear of others. For example, if the child resorts to hostile defenses, he tends to elicit hostility in others. If he becomes severely withdrawn, he tends to elicit withdrawing or rejecting reactions in others. The kinds of negative, defensive behaviors which typically characterize the maladjusted child are precisely the kinds of behaviors which tend to elicit punishing reactions from others in the child's environment. Thus, other persons continue to punish the maladjusted child, and his fear of others persists. Punishment and rejection give rise to fear; fear promotes defensive reactions; and the defensive reactions elicit further punishment. This is the vicious circle of maladjustment. The child lives in an interpersonal world of constant threat. The consequence is unhappiness and fear.

Having started with observations of parent-child interactions, and following, at each step, the implications of research primarily with children, we have arrived at the beginnings, at least, of a theory of children's maladjustment without positing unconscious motives, id and superego conflicts, basic hostilities, unconscious guilts, and the host of intervening variables and hypothetical constructs prevalent in current clinical usage. Nevertheless, the sequence of events we have outlined accounts for the central feature of psychological maladjustment, viz., long-term, persistent fear. This sequence of events can be considered only the beginnings of a theory or an approach to the problem of maladjustment. There are many questions to be investigated before we may profitably build a more detailed and complex theoretical structure. For example, what are the dimensions of emotional sensitivity in the infant, and what are the characteristics of the interaction between the genetically determined sensitivities of the child and the stresses of his interpersonal environment? What are the factors which determine whether or not a child will resort to defenses which elicit threat and punishment from others? Do maladjusted children consistently perceive key figures in their environment as threatening? And do maladjusted children consistently exhibit the kinds of behaviors which tend to elicit threat, rejection, and punishment from others? If the sequence of events we have outlined is a valid description of children's maladjustment, can it also account for maladjustment in adults? Do neurotic adults, for example, perceive their interpersonal environments as threatening, and do they behave in just those ways that tend to reinforce this threat? If the central cause of a person's maladjustment is to be found in his interpersonal relations, what are the psychological implications of various ethical systems and codes of interpersonal behavior?

Our examination of research with children in this attempt to formulate the beginnings of a theory of psychological maladjustment has led to an

interpretation of childhood experience that differs from those theories which started with observations of adults and worked their ways back to childhood by imaginative inference. Notwithstanding the provocative value of inferences about childhood from adult behavior, increased understanding of maladjustment in children depends primarily upon further research focused on the behavior and experience of the child.

References

1. Anderson, H. H., & Anderson, G. L. Social development. In L. Carmichael (Ed.), *Manual of child psychology* (2nd Ed.). New York: Wiley, 1954. Ch. 19.
2. Bishop, B. M. Mother-child interaction and the social behavior of children. *Psychol. Monogr.*, 1951, 65 (Whole No. 328).
3. Brewer, J. E. The measurement of the behavior of second-grade children in relation to the teachers' dominative and socially integrative contacts. Unpublished doctoral dissertation, Univer. of Illinois, 1940.
4. Hattwick, B. W., & Sowell, M. Relation of parental over-attentiveness to children's work habits and social adjustment. *J. educ. Res.*, 1936, 30, 169-176.
5. Jackson, L., & Todd, K. M. *Child treatment and the therapy of play.* (2nd Ed.). London: Methuen, 1948.
6. Johnson, A. M., *et al.* School phobia. *Amer. J. Orthopsychiat.*, 1937, 11, 702-711.
7. Mowrer, O. H. Learning theory and the neurotic paradox. *Amer. J. Orthopsychiat.*, 1948, 18, 571-610.
8. Redl, F., & Wineman, D. *Children who hate.* Glencoe: Free Press, 1951.
9. Sears, R. R. A theoretical framework for personality and social behavior. *Amer. Psychologist,* 1951, 6, 476-483.
10. Shirley, M. M. *The first two years:* Vol. III. *Personality manifestations.* Minneapolis: Univer. of Minnesota Press, 1933.
11. Skeels, H. M., *et al.* A study of environmental stimulation: An orphanage preschool project. *Univer. Iowa Stud. Child Welfare,* 1936, 1, 156.
12. Symonds, P. M. *The dynamics of parent-child relationships.* New York: Bureau of Publications, Teachers Coll., Columbia Univer., 1949.
13. Van Houten, J. Mother-child relationships in twelve cases of school phobia. *Smith Coll. Stud. Soc. Work,* 1947-48, 18, 161-180.

DEVELOPMENTAL DIFFERENCES IN THE STABILITY OF OBJECT PREFERENCES AND CONFLICT BEHAVIOR[1]

BETTY RHEA STEWART
University of Kansas[2]

While children have served as subjects in a number of conflict and choice-making studies, the focus has not been on their performance, but on the testing of general conflict theory. However, the conceptual picture of the young child presented by Baldwin (1), Lewin (5, 6, 7, 8), and Werner (11) suggests that differences in psychological maturity, roughly defined as increasing differentiation in the life space, could be expected to affect the child's stability of object preference and his behavior in a conflict situation.

The young child, as they depict him, lives in a here-and-now world where perception, emotion, action, and general cognitive activities are still partially fused. As the child grows older, differentiation between and within these capacities develops. He has an extended range of emotions instead of his earlier diffuse reactions. His perceptions are less affectively tinged and are well articulated rather than global. He has a past and a future which are significant influences in his present. He is able to withhold action until he has grasped a situation cognitively without being forced to act by the pull of the immediate situation.

The study was designed to test the effects of increasing age upon stability of object preferences and upon behavior in a conflict situation of the approach-approach type. The following hypotheses, based on the theoretical framework offered by Baldwin, Lewin, and Werner, were made about differences in the behavior of the young preschool age child and the older child.

1. The young child's preferences for objects will be less stable than the older child's, both in an immediate situation and after an interval of time.

As Werner portrays the young child, his behavior is governed largely by his affective needs. These needs, according to Lewin (5), shift often, causing alterations in the attractiveness of objects in the child's environ-

[1] This study was performed as part of a dissertation submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy in the Department of Psychology, University of Kansas. The writer is indebted to Dr. Edward L. Wike for his direction of the research and to Dr. Anthony J. Smith and Dr. Roger Barker for their assistance.

[2] Now at Department of Child Study, Vassar College.

ment. Therefore, in expressing preference for one object rather than another, the young child may change his choice within a comparatively short time without realizing the inconsistency of his action. Even greater disparity may be predicted if the time interval between choices is long, since his time perspective is limited.

The older child, in contrast, should be more consistent in his choices. According to Lewinian theory, the psychological world of the child becomes more stable and more differentiated with increasing age, so the older child should be less at the mercy of his immediate needs and the environmental forces of the moment. In addition, his time perspective is increasing, thus making it possible for his past actions to have implications for his present behavior.

2. Young children will agree less among themselves on their preferences for objects than will older children.

Lewin lists as one of the principal differences in developmental stages "an increasing *differentiation* of every level of the life space into a multitude of social relations and areas of activities" (7, p. 797). The young child does not have a clearly defined age and sex role. He is still at liberty to express a large range of preferences which do not have to conform to those of other children in his group. The older child's preferences, however, are more often imposed on him by the expectations of adults and his peers. A boy, for example, will incur ridicule if he displays interest in dolls and a girl will be criticized if she is too active in boys' games.

3. In an approach-approach conflict situation the young child will make a decision more quickly, but less decisively, than the older child.

Werner describes the young child's world as one of action, where perception and direct motor activity often function as a single unit. The young child, seeing two desirable objects side by side, can be expected, according to this view, to react to them with some kind of direct motor behavior. But both Werner and Lewin indicate that while the young child may act impulsively in a choice situation, he will not necessarily act decisively.

The older child, with his greater self-control in the face of internal and external pressures, should take longer to make his choices, and should make them with more certainty.

4. In an approach-approach conflict situation the young child's choices will be less influenced by differing valences of objects, as ranked by himself, than will the older child's choices.

The young child may experience the presentation of each set of choice objects as a total situation and not as related to previous choices from a larger group of objects. The young child's world has not yet expanded logically, according to Baldwin, so it is difficult for him to perceive how isolated aspects of a situation fit together to form a total pattern.

The older child, with his longer time perspective and his greater ability to comprehend interrelationships, may be influenced in his present choices by his earlier statements of preference for the objects.

Method

Subjects

The younger group was composed of 30 children from the University of Kansas nursery school, ranging from 3 years to 5 years, 4 months, with an average age of 4 years. There were 16 girls and 14 boys in the sample. As a group they were intellectually superior with a mean IQ of 131 on the Stanford-Binet test, Form M, and a range from 108 to 152.

The older group was made up of 30 fourth graders from a Lawrence, Kansas, public school. Their average age was 9 years, 10 months, with a range from 9 years, 3 months, to 10 years, 4 months. There were 14 girls and 16 boys in the group. Their mean IQ, based on a combination of the California Mental Maturity Tests for elementary grades, and the Stanford-Binet, Form L, was 124, with a low score of 103 and a high score of 151.

Materials

Materials used in the experiment were 12 ten cent toys: doll, comb and mirror set, jacks, marbles, ray gun, tank, pencil, paper pad, paints, telescope, puzzles, and toy snake.

Procedure

Ranking procedure. Pilot study results indicated that young children cannot rank toys in the usual way. The present method gave a system of preferential ranking of the toys which the younger group could understand and perform.

Each child was shown a 20 by 15 inch box with a hinged lid, containing all the toys arranged in a standard order. Although the child was not allowed to become involved in a game with the toys, his attention was called to all of them, and he was encouraged to explore their possibilities. When he had examined the toys, he was told to look at *all* the toys carefully and to pick out the one toy he liked best. The child's choice was timed with a stop watch from the moment the instructions were completed until he made his choice verbally, by pointing, or by picking up the toy. The toy was then removed from view and the lid of the box was closed until the child was again told to choose the toy he liked best of the ones remaining. The procedure was repeated and times were recorded until all the toys were chosen.

Paired comparisons procedure. Twelve combinations of toys were arranged from the three toys the child had ranked as his top choices and the three toys that he had liked least. The child was presented either with a pair of toys he had ranked high or a pair he had ranked low; he was never given a pair made up of one high and one low ranking toy. The order of presentation of the paired comparisons was randomized so that no subjects in a sample received the toys in exactly the same order. Each toy appeared once on the right and once on the left. In addition, the order

of appearance for most and least liked pairs of toys was randomized for each subject.

Immediately after the ranking procedure a 20 by 20 inch cardboard screen was placed on the table between *S* and *E*. The child was told that two toys would be placed behind the screen. When *E* lifted the screen, *S* was to look at both toys carefully and to indicate the toy he liked "the very best." The screen was placed about six inches from the edge of the table and the toys were put just behind the screen approximately 12 inches apart. Choices were timed with a stop watch and choice time was defined as the period of time extending from the instant the screen was lifted until a final choice was indicated. But times were also noted if *S* grabbed both toys or pointed quickly from one toy to the other before making his final choice.

Ranking procedure retest. From 8 to 11 days later subjects in both groups were again asked to rank the toys, the same procedure being used as in the first ranking situation. This completed the experiment with the younger children. The older children were interviewed briefly about their general reactions to the toys.

Results

Mean scores, ordinarily the appropriate measures of central tendency for such data as the individual ranking times and paired comparisons choice times, would have been misleading, since the raw data for many of the subjects were markedly skewed. Median ranking times and choice times for each subject were employed throughout the study as the best representatives of individual scores. Rank order correlations between the first and second ranking orders, computed for each subject, were treated as indices of stability of preference and were treated as individual scores in subsequent analyses. Because the frequency distributions of median choice times and of the rank order correlations were also skewed, the results throughout were analyzed by nonparametric statistics.

Both groups were significantly consistent in choices from first to second rankings, as indicated by a mean rank order correlation of .41 for the nursery school group and .80 for the grade school group. However, in the frequency distributions of rhos for the two groups, the young children's scores fell into a roughly bimodal distribution, with some children evidencing little stability of choice from one ranking to the other, while others were highly consistent. The grade school children's scores were concentrated at the upper end of the distribution. Wilcoxon's *T* test for unpaired replicates, applied to rhos of the two groups, yielded a critical ratio of 4.91, significant at far less than the .001 level. This finding strongly supported the hypothesis that the older children would show greater stability of object preferences over an extended period of time than the younger children.

To test the stability of choice behavior for young and older children when only a short period of time elapsed, comparisons were made between preferences expressed on the first ranking and those indicated immediately afterward in the paired comparisons procedure. The data were categorized as highly consistent, consistent, and inconsistent. If, in the paired comparisons situation, the child chose in all or all but two of his 12 choices the toy he had ranked higher a few minutes before, he was considered highly consistent. If he chose the toy with the higher ranking between half and three-fourths of the time, he was categorized as consistent. His behavior was scored inconsistent if he chose a toy he had ranked lower more than half of the time. A chi square of 14.72 (2 *df*), significant at less than the .001 level, indicated that nursery school children were much less consistent in their choices than grade school children when a short period of time separated expressions of preference.

It was hypothesized that members of the younger group would differ widely from one another in their rankings of the 12 toys, while the preferential rankings of the older children were expected to be quite consistent from individual to individual. To test this assumption, Kendall's coefficient of concordance (*W*) was computed for each group's first order of ranking and on the data for each group broken down according to sex. The significance of the *W*s was tested by transforming them into *F*s and evaluating the *F* distribution according to a method described by Walker and Lev (10). Results of the analysis are presented in Table 1.

It is evident from the table that the fourth grade group agreed only a little better among themselves on their preferences than the younger children. But, when the data were analyzed according to sex for the two groups, the seeming lack of agreement among the older group appeared to be an artifact caused by sex differences in toy preferences. Since the grade school boys were displaying a consistent order of preference for certain toys while the girls were showing a consistent order for different toys, the real stability of the preferences was concealed by the statistical

TABLE 1

COEFFICIENTS OF CONCORDANCE FOR TOY PREFERENCES

Group	*W*	*df*	*F*
Nursery School Total	.205	330	7.48*
Grade School Total	.233	330	8.79*
Nursery School Boys	.301	154	5.60*
Nursery School Girls	.215	176	4.11*
Grade School Boys	.599	176	22.41*
Grade School Girls	.465	154	11.30*

* Significant at .001 level.

method. There were also indications that the boys in both groups agreed more in expressions of preference than the girls, though this cannot be stated conclusively, since there is no significant test of differences between *W*'s available.

A comparison of ranking times for the nursery school group and the fourth graders disclosed that the younger children chose more rapidly than the older children when confronted with the whole array of toys in the first ranking situation. Differences according to Wilcoxon's *T* were significant at the .01 level, the nursery school group having a median of 2.42 seconds and the grade school children a median of 3.16 seconds. Differences between the groups were not significant on second ranking times.

When the total choice times for the two groups were compared for the paired comparisons situation, differences between the groups were significant at the .001 level, the nursery school group taking much longer to make their choices than the grade school children. Median time for the nursery school choices was 3.76 seconds, while the grade school children had a median time of 2.07 seconds. This finding was in direct contrast to the original hypothesis that the young children would react faster than the older ones in a choice situation.

TABLE 2

COMPARISON OF CHOICE TIMES FOR NURSERY SCHOOL CHILDREN AND GRADE SCHOOL CHILDREN ON MOST LIKED TOYS AND LEAST LIKED TOYS

Group	*Valence of Toys*	*Median Choice Time*	*Difference*
		WITHIN GROUPS	
Nursery School	Most Liked	3.76	.12
	Least Liked	3.64	
Grade School	Most Liked	1.94	.15*
	Least Liked	2.09	
		BETWEEN GROUPS	
Nursery School	Most Liked	3.76	1.82†
Grade School	Most Liked	1.94	
		BETWEEN GROUPS	
Nursery School	Least Liked	3.64	1.55†
Grade School	Least Liked	2.09	

* Significant at .002 level.
† Significant at .001 level.

The second part of the hypothesis stated that young children would be more indecisive than older ones. Choice behavior was considered indecisive (a) if the child grabbed both toys and had to be reminded that he could choose only one, (b) if he verbally expressed a desire to have both toys and did not choose one, or (c) if he wavered between the two toys, first indicating one and then the other. Seven of the 15 nursery school children displaying indecisive behavior grabbed both toys, while none of the four grade school children categorized as indecisive did this. All indecisive behavior of the grade school children was shown, instead, by a wavering between the two choice objects, with no verbally expressed desires for both. Four of the nursery school children, in contrast, said they wanted both, and four others indicated one toy, then the other in their choices. A chi square of 11.14 (1 *df*), significant at the .001 level, supported the hypothesis that the younger children would be more indecisive.

As Table 2 indicates, the younger subjects chose toys in the paired comparisons situation as rapidly when they had ranked them high as when they had ranked them low, a sign test computed on the difference between median choice times for the most and least liked toys indicating nonsignificance. The older children, on the other hand, took significantly longer to choose toys they had ranked low than those they had ranked high, a sign test indicating differences significant at the .002 level.

When comparisons were made between the two groups rather than within each group the difference between median choice times on the most liked toys was significant at the .001 level as tested by Wilcoxon's *T*, with the nursery school children taking longer. The nursery school children were also slower than the older group in choices between toys they liked least, the difference again being significant at the .001 level.

Discussion

The design of the experiment corresponded closely to an approach-approach conflict situation, since, theoretically, choices were to be made between objects of varying degrees of positive valence or between objects about which the subjects did not feel strongly. None of the choices was intended to cause active avoidance-avoidance reactions, as is the case, for example, in animal experiments where shock is utilized. However, the behavior of some of the older children indicated that valences of least liked toys were negative enough at times to represent a mild avoidance-avoidance conflict.

The choice situation for the younger children more nearly resembled an approach-approach situation for all the toys. However, if the approach-approach quality of the conflict situation were the only factor operating, the young children's choice times for *all* the toys should have been approximately the same as for the older group's choice times on the *most liked* toys, an approach-approach situation also. According to conflict theory,

the longest choice times for either group should have been the older group's time on the least liked toys, an avoidance-avoidance conflict. But the nursery school children were significantly slower in all their choices, which would indicate that the level of psychological maturity must be taken into account also in making predictions about children's conflict behavior.

The hypothesis that young children would react faster than older children to the choice materials was the only major prediction not supported by the data. Studies have shown that young children's reaction times are slower than older children and adults (3), but observations of the nursery school group's choice making behavior indicated that the process of deciding was in itself difficult for them. Subjects who grabbed both toys often had to be told more than once that they could choose only one toy and even then they had difficulty reaching a final decision. Contrary to Lewin's theory of conflict, which states that the region of indecision is less attractive when the conflict is between objects of positive valence, the young children seemed to find it more satisfying to remain in an indecisive state than to give up definitely the possibility of having both toys. Barker (2) suggests that in an approach-approach conflict, where the valences are both highly positive and approximately equal, a relatively greater increase in the strength of one force will be necessary to resolve the conflict than where the two objects are intrinsically less attractive, but still positive. The intensity of the young children's conflicts in choosing between two liked toys may have been stronger than the intensity of the older children's conflicts to the extent that the toys were intrinsically more attractive for the younger than the older children.

That young children in general may have a more undifferentiated reaction to valence was suggested by their inability to rank in the usual way. The concept of an order of preference, where over-all feeling of liking or disliking an object are secondary to the relationship between the attractiveness of one toy and another, appeared to be a very difficult idea for the young child to grasp. In several instances the nursery school children had difficulty making selections even in the modified ranking situation because they could not conceive of liking any single toy best in a group where they disliked all the toys from which the choice was to be made.

Summary and Conclusions

The experiment was designed to examine the effects of increasing age upon stability of object preferences and behavior in conflict situations of an approach-approach type, chronological age being the independent variable. Psychological maturity and valence were assumed to be intervening variables. The dependent variables were ranking times, choice times between paired comparisons of the toys, and stability of preferences over periods of a few minutes and approximately a week.

The two samples consisted of 30 nursery school children and 30 fourth grade school children. In the first part of the experiment subjects were shown 12 ten cent toys which they ranked by a modified ranking procedure. In the second phase the children chose between pairs of most liked and least liked toys as determined by their previous rankings. The third part consisted of a retest of ranking behavior after an 8 to 11 day lapse, and a brief interview with the older children. The following conclusions were drawn from the results:

1. Young children are less consistent in their choices than older children both in an immediate choice situation and after an interval of time.

2. Young children differ more from each other in their preferences than do grade school children.

3. Young children choose more rapidly than older children when given a whole array of toys to choose from, but less rapidly when the task is to choose one of two toys.

4. The behavior of young children in a choice situation is more indecisive than older children's behavior.

5. The choice behavior of young children is not influenced by the valence of the objects, as ranked by themselves, but older children find it more difficult to make choices between objects they dislike.

References

1. Baldwin, A. L. *Behavior and development in childhood.* New York: Dryden, 1955.

2. Barker, R. G. An experimental study of conflict by children; time elapsing and amount of vicarious trial and error behavior occurring. In Q. McNemar & M. A. Merrill (Eds.), *Studies of personality.* New York: McGraw-Hill, 1942.

3. Jones, H. E. Reaction time and motor development. *Amer. J. Psychol.,* 1937, 50, 181-194.

4. Kendall, M. G. *Rank correlation methods.* London: Charles Griffin, 1948.

5. Lewin, K. Environmental forces in child behavior and development. In Carl Murchison (Ed.), *A handbook of child psychology.* Worcester, Mass.: Clark Univer. Press, 1931.

6. Lewin, K. The conceptual representation and measurement of psychological forces. In Donald K. Adams & Helge Lundholm (Eds.), *Contributions to psychological theory.* Durham, N.C.: Duke Univer. Press, 1938.

7. Lewin, K. Behavior and development as a function of the total situation. In Leonard Carmichael (Ed.), *Manual of child psychology.* Worcester, Mass.: Clark Univer. Press, 1946.

8. LEWIN, K. Regression, retrogression, and development. In Dorwin Cartwright (Ed.), *Field theory in social science.* New York: Harper, 1951.

9. MILLER, N. E. Experimental studies in conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders,* Vol. I. New York: Ronald Press, 1944.

10. WALKER, HELEN M., & LEV, J. *Statistical inference.* New York: Henry Holt, 1953.

11. WERNER, H. *Comparative psychology of mental development.* New York: Follett, 1948.

12. WILCOXON, F. *Some rapid approximate statistical procedures.* New York: American Cyanamid Company, 1949.

A TIME SERIES ANALYSIS OF GROWTH IN NOSE HEIGHT DURING CHILDHOOD[1]

HOWARD V. MEREDITH
Iowa Child Welfare Research Station, State University of Iowa

This study pertains to the growth of the individual child. It is related historically to the pioneering longitudinal studies of Montbeillard (23) in the eighteenth century, Bowditch (4) in the nineteenth century, and Baldwin (3) in the early years of the present century. At closer range, it is a part of a program of research on individual growth fostered at the State University of Iowa during the past two decades (16).

In the pursuit of biologically valid knowledge on individual growth it is important to select for study measures that can be obtained with very high dependability (16); skeletal nose height is such a measure. Seriatim data for this dimension, spaced at yearly intervals, have been amassed and analyzed.

The findings are organized under three headings: (1) trend, (2) velocity, and (3) relation with other variables. Primary emphasis is placed on the classification of magnitude trends, the analysis of velocity differences, and the comparison of body traits with respect to growth rate correlations and increment trends.[2]

SUBJECTS

The subjects were 80 children, 38 boys and 42 girls, residing in or near Iowa City, Iowa. All were voluntary participants in the Facial Growth Study, a long-term research program at the State University of Iowa begun in 1946 under the joint sponsorship of the Department of Orthodontics and the Iowa Child Welfare Research Station. Enrollment in the study was based on likelihood of continuing residence in the community and

[1] Supported in part by a research grant, D-217, from the National Institute for Dental Research, of the National Institutes of Health, U.S. Public Health Service.

[2] Using terminology appropriate to the view that growth is "a form of motion," Tanner (26) refers to magnitude or integral trends as "distance" curves, and to increment or differential trends as "velocity" curves.

willingness to participate; it was not related to odontic condition or orthodontic need.

The sample comprised a group of American-born white children predominantly of northwest European ancestry and above average in socioeconomic status. Every subject was examined annually from age 5 years to age 12 years. Forty-five of the subjects, 19 boys and 26 girls, were examined also at ages 13 years and 14 years. The examinations were made on or near each child's birthday anniversary.

Data

Data for skeletal nose height were derived from *norma lateralis* roentgenograms of the face. The roentgenograms were obtained with (a) the head maintained in a Higley (9) positioner, (b) the film paralleling the saggital plane near the side of the head, (c) the distance from the X-ray source to the saggital plane held constant at 150 cm., and (d) the central ray passing through the acoustic meatuses.

To minimize age-to-age variation in landmark determination, a given landmark was registered at one sitting on the entire series of roentgenograms for a subject. For example, two persons experienced in roentgenographic anthropometry examined the nasion region of the films for Subject F1 and reached agreement on the location of nasion. This was done with the aid of a magnifying glass, and the point decided upon for each film was pricked with a needle probe. Employing the same procedure, the most anterior point of the anterior nasal spine was marked.

The roentgenographic distance from the superior landmark (nasion) to the inferior landmark (tip of anterior nasal spine) was measured by means of a steel millimeter tape read to the nearest 0.01 cm. Films were taken for measurement in random order, and each was measured by two anthropometrists working independently. When the two independent records were found to agree within 0.02 cm., they were averaged. In instances of greater discrepancy (there were less than 10 per cent of these) two additional measurements were taken and the mean of the four determined.

These measures derived from the films were corrected for roentgenographic enlargement and blurring. A previous publication (21) describes in detail the methods employed to obtain the corrections. It will suffice here to record that the roentgenographic measures were converted to measures of skeletal nose height by subtraction of amounts varying from 0.23 cm. to 0.41 cm.

Summarily, meticulous care was taken to obtain valid measures of skeletal nose height. Attention was given to both the reliability of the data (by controls to minimize chance error in subject orientation, landmark location, and measurement) and their relevance (by corrections for the systematic errors that roentgenography entailed).

Findings for Trend

The Group Trend

Table 1 presents means for skeletal nose height derived from the measures on those subjects studied annually from age 5 years to age 12 years. It will be seen that the means for boys and girls (a) do not differ at any age by more than 0.06 cm. and (b) increase with age in a highly similar manner. In view of these findings, attention will be turned to the central tendency trend for the two sexes combined.

Table 1

ARITHMETIC MEANS FOR SKELETAL NOSE HEIGHT (NASION TO ANTERIOR NASAL SPINE IN CENTIMETERS) ON 80 NORTH AMERICAN WHITE CHILDREN STUDIED LONGITUDINALLY

Age (yrs.)	*Boys* ($N = 38$)	*Girls* ($N = 42$)	*Both Sexes* ($N = 80$)
5	3.95*	3.89	3.92
6	4.11	4.06	4.08
7	4.25	4.20	4.22
8	4.37	4.33	4.35
9	4.49	4.46	4.48
10	4.60	4.57	4.59
11	4.70	4.68	4.69
12	4.79	4.77	4.78

* The reader may wish to interpret the sample means for a particular age as parameter estimates. Standard errors of the mean are 0.03 cm. for boys 5 years of age, 0.02 cm. for girls 5 years of age, and 0.03 cm. to 0.04 cm. in all other instances.

Figure 1 shows that the curve drawn to the composite means for both sexes from Table 1 has an upward slope of diminishing steepness, i.e., the group magnitude trend for skeletal nose height over the childhood period between ages 5 years and 12 years is a negatively accelerated increasing function of age. An equation for this trend, obtained by the method of least squares, is: Skeletal Nose Height $= 2.710\ (\text{Age})^{0.2285}$. Goodness of fit is satisfactory; the difference between comparable means, estimated and obtained, in no instance exceeds 0.01 cm.

Central tendency trends of the same general form (increase during childhood at a declining rate) have been obtained for bizygomatic width of the upper face (14) and bigonial width of the lower face (21). The former study extends over the childhood period 4 to 10 years and the latter over the period 5 to 11 years.

To this point, the group trend prior to age 12 years has been considered. Although the size of the sample drops from 80 to 45 after age 12 years,

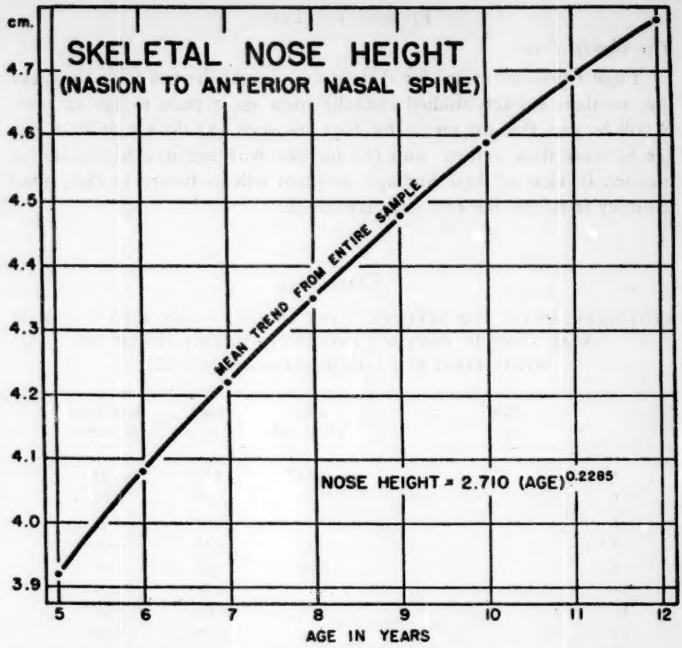

FIGURE 1—Magnitude curve drawn to group means for 80 North American white children of both sexes studied longitudinally.

some inquiry up to age 14 years is possible. Means on this smaller sample were obtained at ages 12, 13, and 14 years. The mean at age 12 years is 4.78 cm., the same as that for the full sample of 80 children (*see* Table 1). At the two succeeding ages the means are 4.88 cm. and 4.97 cm., respectively. These figures, together with the values from Table 1 for ages 10, 11, and 12 years, warrant the inference that between ages 10 and 14 years the group trend for skeletal nose height approximates a slowly rising rectilinear function of age. A similar group trend for this age span has been found from numerous investigations on width of the upper face (14).

The means of Table 1, in common with those reported in a few other investigations on both living and non-living children (1, 12), are for skeletal nose height. Another group of studies on the living child reports means based on direct measurement of the nose from the level of nasion to subnasale, i.e., means for integumental nose height (5, 6, 7, 8). To convert the values from one approach to equivalent values for the other approach necessitates information on the comparability of results when the two

methods are employed on the same group of subjects. Measures of integumental nose height at age 5 years, obtained using the Ashley-Montagu (2) procedure for locating the level of nasion, were available for the 80 subjects of the present study. The mean of these measures (4.13 cm.) exceeds its corresponding value for skeletal nose height by 0.21 cm.

Returning to findings specific for sex, although the Table 1 means on boys are slightly higher than those on girls the sex difference is not statistically significant. Nor can it be posited on the basis of previous studies that there is a sex difference in nose height during the childhood years (1, 5, 7, 25, 28). However, the availability of additional subjects at age 5 years made it possible to test for sex differences in skeletal nose height on samples of 82 boys and 84 girls. These enlarged samples gave means of 3.98 cm. and 3.86 cm. for boys and girls, respectively. The difference is significant at a high level of confidence ($t = 3.58$, df 164), and harmonizes with a substantial body of knowledge for other facial dimensions showing that on the average boys are larger than girls. Variability findings from these enlarged samples at age 5 years are as follows: boys, standard deviation 0.23 cm. and coefficient of variation 5.7 per cent; girls, standard deviation 0.20 cm. and coefficient of variation 5.2 per cent. The obtained differences in variability are nonsignificant statistically.

Trends for the Individual

The procedure used in analyzing individual magnitude trends for skeletal nose height was as follows:

1. The seriatim data from each subject was plotted on a Cartesian grid. As is customary, ages were used as abscissa values and measures of skeletal nose height as ordinate values.

2. A curve was drawn connecting each series of plotted points, yielding a total of 80 curves. Individually, these curves showed clear expression of trend, i.e., there was little manifestation of random or cyclic fluctuation (16, 21, 24). Utilizing a phrase from Quenouille (22), the time series on each child consisted "largely of a trend."

3. The curves were classified with respect to similarity of trend. That is, rising rectilinear trends were grouped in one class, trends that ascended in a path concave to the age base line in another class, and so forth.

4. A composite trend representing each grouping of individual curves was derived. This was done by tabulating the measures on each member of a class according to age, computing means for the class at successive yearly ages, plotting these class means, and drawing a curve to them.

Findings will be presented first for the larger sample studied between ages 5 and 12 years, then for the smaller sample studied between ages 5 and 14 years.

For the period from age 5 years to age 12 years, three types of individual magnitude trend were found for skeletal nose height. The commonest type is an ascending curvilinear trend concave downward; this

represents more than two-thirds of the 80 subjects studied. The other types are: an ascending rectilinear trend, and an ascending curvilinear trend concave downward in the first half of the period and convex downward in the last half. Figure 2 portrays the composite trends derived to delineate the three types, and indicates the absolute and relative frequencies of each type in the sample studied.

The lower right panel of Figure 2 displays two individual trends illustrating the range of variation within the concave class. Although both

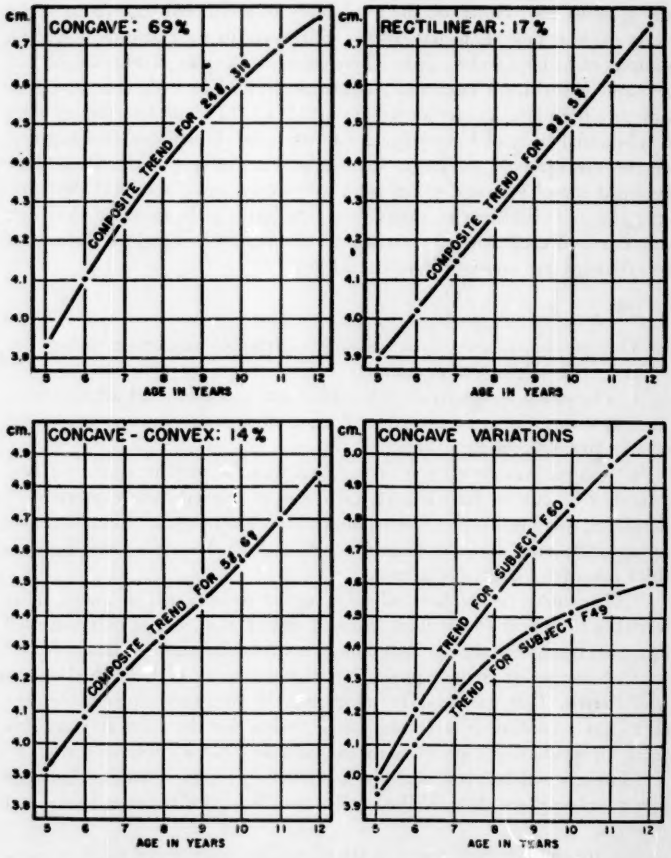

FIGURE 2—Magnitude curves portraying varieties of individual growth in skeletal nose height during the childhood period between 5 and 12 years of age.

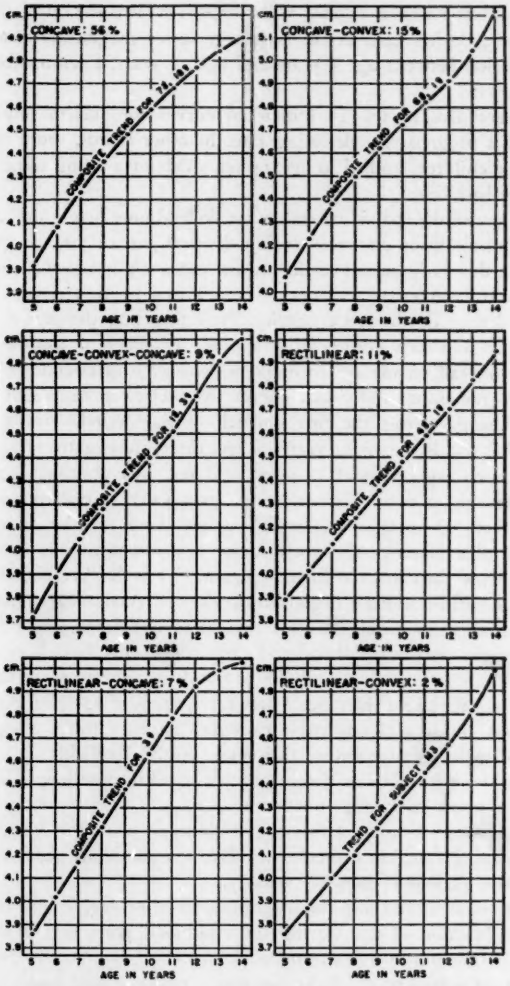

FIGURE 3—Magnitude curves depicting types of growth in skeletal nose height between 5 and 14 years of age.

trends manifest an upward course of declining steepness, the over-all rise is 0.41 cm. more for Subject F60 than for F49. Further consideration of individual differences in velocity will comprise a later section of the paper.

Figure 3 depicts each of the six categories of magnitude trend found for the period between age 5 years and age 14 years. Examination of this graph shows:

1. Twenty-five of the 45 individual curves for skeletal nose height ascend at a progressively declining rate. In other words, more than one-half of the children studied to age 14 years exhibit a rising trend concave to the age base line.

2. Eleven of the individual curves for skeletal nose height are ascending trends of diminishing steepness during the first part of the age span only. For seven, this early period of downward concavity is followed by a span of downward convexity; for four, it is followed by first downward convexity and then downward concavity. Projecting into the period between age 14 years and early adulthood, it appears reasonable to assume that the "concave-convex" trends are incomplete "concave-convex-concave" trends.

3. The remaining individual curves for skeletal nose height are rectilinear during at least the first one-half of the period. Five are rising straight-line trends out to 14 years of age, three are negatively accelerated trends between ages 12 and 14 years, and one is a positively accelerated trend after age 12 years. Extrapolating again (16), it may reasonably be inferred that (a) to the "rectilinear" trends there will be added either a segment that is concave downward, or a segment that is first convex downward and then concave downward, and (b) to the "rectilinear-convex" trend there will be added a negatively accelerated (concave) segment.

Synthesizing, magnitude trends representing individual growth in skeletal nose height during childhood and early adolescence have been graphed (Figures 2 and 3) and discussed. These trends in part support and in part suggest the generalized inference: Between 5 years of age and early adulthood, white children of northwest European ancestry travel along one of four trend paths in augmenting their nose height. Explicitly,

1. For about 1 child in 2 the magnitude trend of nose height is a negatively accelerated increasing function of age.

2. For about 1 child in 4 the magnitude trend of nose height is a two-cycle increasing function of age, i.e., there is a period of positive velocity between two periods of negative velocity.

3. For about 1 child in 8 the magnitude trend of nose height is a rising rectilinear function of age during the childhood years and a negatively accelerated function of age during adolescence.

4. For an occasional child the magnitude trend of nose height proceeds at a constant absolute rate during childhood, at an increasing absolute rate in early adolescence, and at a declining absolute rate in late adolescence.

It should be recognized that the foregoing findings stem from the Cartesian grid approach to the study of individual growth. From the semilog-

arithmic grid approach, which also is useful in biometric research, somewhat different trend groupings would be expected (21).

Findings for Velocity

Individual and Age Differences

Variability of velocity in skeletal nose height was investigated on the 80 subjects studied from age 5 years to age 12 years. Centimeter and percentage increments were determined for four annual intervals and the full seven-year period. Each series of derived values was arranged in order of magnitude. Statistical reductions for these 10 ordered series are assembled in Table 2. Selected findings are:

1. There is a general decrease in velocity of growth in nose height with age. In centimeter terms, the minimum increase during the sixth postnatal year coincides with the 30th percentile of the increment distribution for the twelfth year; the 70th percentile of the distribution for the fifth year coincides with the maximum increase during the twelfth year; and so forth. In percentage terms, the typical gain for the sixth year (4.0 per cent) is twice that for the twelfth year, and the terminal values of the annual percentage gain distributions decline from 2.1–6.1 to 0.7–3.9.

2. Over the seven-year interval between age 5 years and age 12 years, increase in nose height for some children is double that for other children.

Table 2

VARIABILITY OF INCREASE IN SKELETAL NOSE HEIGHT FOR FOUR ANNUAL PERIODS AND A SEVEN-YEAR PERIOD ON 80 NORTH AMERICAN WHITE CHILDREN OF BOTH SEXES

Age Interval (yrs.)	*Minimum*	PERCENTILES: 10	30	70	90	*Maximum*
		CENTIMETER INCREASE				
5 – 6	0.08	0.12	0.14	0.17	0.20	0.24
7 – 8	0.06	0.10	0.11	0.14	0.16	0.19
9 – 10	0.05	0.08	0.10	0.12	0.14	0.17
11 – 12	0.03	0.05	0.08	0.11	0.14	0.17
5 – 12	0.55	0.71	0.79	0.91	1.01	1.15
		PERCENTAGE INCREASE*				
5 – 6	2.1	3.0	3.6	4.4	5.0	6.1
7 – 8	1.3	2.3	2.7	3.3	3.8	4.5
9 – 10	1.1	1.7	2.2	2.8	3.2	4.0
11 – 12	0.7	1.1	1.7	2.4	2.9	3.9
5 – 12	14.5	17.7	19.8	23.8	25.9	30.2

* Computed as: Absolute increase in skeletal nose height during the age interval indicated × 100 / Magnitude of skeletal nose height at the age the interval begins.

The obtained absolute and relative dispersions extend from 0.55 cm. to 1.15 cm., and from 14.5 per cent to 30.2 per cent. For the absolute increments as a series, 1 in 10 lies below 0.71 cm., 1 in 4 between 0.79 cm. and 0.91 cm., and 1 in 10 above 1.01 cm. Corresponding subdivision values for the percentage increment series are 17.7, 19.8, 23.8 and 25.9, respectively.

Correlation Findings

Two varieties of relationship involving rate of growth in skeletal nose height were studied; they may be designated (a) size-gain and (b) gain-gain. The Pearson product-moment method r was used; all analyses were based on data for 80 subjects. Size-gain findings are:

1. The correlation is close to zero ($r = -0.04$) for the association between nose height at age 5 years and centimeter increase in nose height over the age span 5 to 12 years.

2. There is low negative correlation ($r = -0.35$, df 78, $t = 3.72$) denoting the association between nose height at age 5 years and percentage increase in nose height over the age span 5 to 12 years.

Several other studies have reported size-gain correlation coefficients for facial dimensions. These studies encompass: bizygomatic width of the upper face, size at 4 years with gain from 4 to 9 years (14); bigonial width of the lower face, size at 5 years with gain from 5 to 11 years (21); dental arch width at the deciduous canines, size at 4 years with gain from 4 to 8 years (10); and dental arch width at the deciduous second molars, size at 4 years with gain from 4 to 8 years (18). In all instances the obtained relationships were low, statistically nonsignificant, or zero.

The second variety of velocity association investigated for skeletal nose height was the correlation between increase from 5 to 7 years of age and increase from 7 to 12 years of age. An answer was sought to the question, Can rate of growth in late childhood be predicted efficiently from rate of growth in early childhood? Again, analyses were made for both absolute and relative gains; the results are $r = 0.18$ and $r = 0.15$, respectively. In neither case can the null hypothesis be rejected; predictive possibilities are negligibly poor.

Findings on Relation with Other Variables

Velocity Correlations

Associations were investigated between gain in skeletal nose height during the period from 5 to 10 years of age, and gain in each of seven other dimensions over the same period. Six of the dimensions, together with references that describe the procedure employed in their measurement, are: head length (11), head width (13), upper face width (14), lower face width (21), hip width (17), and stature (19). The remaining dimension, subnasal height, was determined as the distance from the most anterior projection of the anterior nasal spine to menton. As in the previous section,

employing the Pearson r formula, both absolute and relative increments were correlated. Table 3 displays the coefficients obtained. Selected findings are:

1. It is not possible to reject the hypothesis of zero correlation between increase in nose height and increase in that portion of face height below the nose.[3] Also, null hypotheses cannot be rejected for gain in nose height with gain in bigonial face width, and gain in bi-iliocristal hip width.

2. The hypothesis that there is no relationship between increase in nose height and increase in head length can be rejected at the 5 per cent level of confidence. For gain in nose height with gain in stature, the null hypothesis can be rejected at the 1 per cent confidence level.

3. All of the increment relationships studied for the childhood span from 5 to 10 years of age are weak. This is shown by correlation coefficients ranging from low positive to zero.

TABLE 3

INCREASE IN SKELETAL NOSE HEIGHT BETWEEN 5 AND 10 YEARS OF AGE (VARIABLE *X*) CORRELATED WITH INCREASE IN SEVEN OTHER VARIABLES OVER THE SAME AGE PERIOD (VARIABLE *Y*)

Variable Y	*N*	r*	p	r†	p
Head length	80	0.27	0.05	0.23	0.05
Head width	80	0.24	0.05	0.18	...
Subnasal height	66‡	—0.05	...	—0.05	...
Bigonial face width	80	0.12	...	0.12	...
Bizygomatic face width	80	0.25	0.05	0.19	...
Bi-iliocristal hip width	80	0.17	...	0.21	...
Stature	80	0.34	0.01	0.35	0.01

* Coefficients for absolute (centimeter) increments.
† Coefficients for relative (percentage) increments.
‡ Increments for subnasal height were obtained only from roentgenograms showing occlusion of the upper and lower teeth at both ages.

Velocity Trend Comparisons

It has been stated in recent publications that velocity curves for facial dimensions show "a general circumpuberal rise" (20) and, with more explicit reference to the subject matter of this paper, that there is "a distinct acceleration of nose growth" at adolescence (27). The applicability of these statements to growth in skeletal nose height cannot be fully evaluated with data extending only to 14 years of age. There is the possibility, however, of supplementing the findings of the earlier section on magnitude trends

[3] At age 5 years, roentgenograms satisfactory for deriving measures of the nasal and subnasal components of skeletal face height were available on 138 children. The coefficient obtained from correlating these measures is $r = 0.17$.

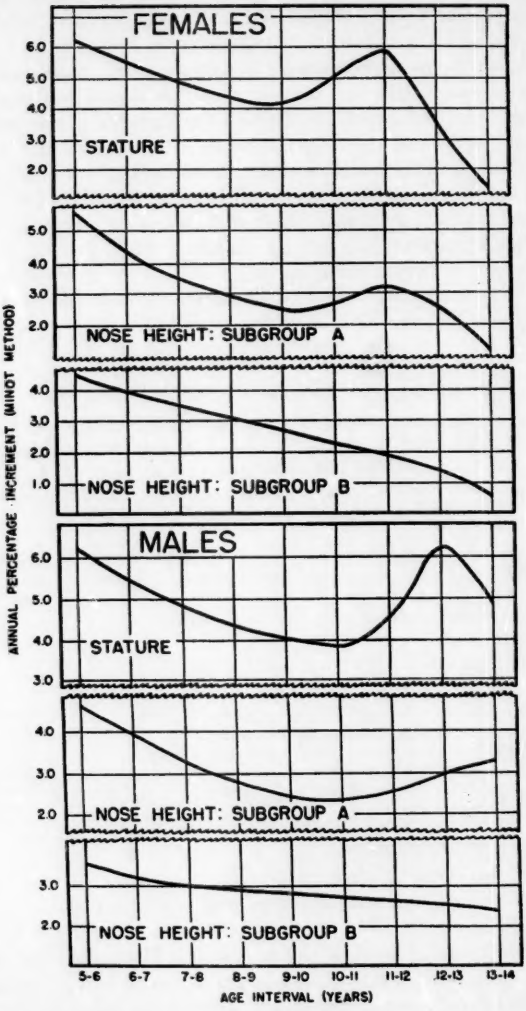

FIGURE 4—Smoothed velocity curves for selected stature groups and skeletal nose height subgroups (*see* text).

with some findings from comparing velocity trends for nose height and stature.

From among the 45 subjects on whom data were available to age 14 years, it was possible to draw two small groups that were homogeneous with respect to age at which the crest of the circumpuberal acceleration in stature was attained. One group consisted of seven girls cresting between 11 and 12 years; the other of seven boys cresting between 12 and 13 years.

On each member of both groups, annual percentage increments were computed for stature and skeletal nose height. The percentage increments for stature were tabulated in age-sex classes. Their subsequent analysis is indicated by stating that the first and fourth velocity trends shown in Figure 4 are drawn to means for these classes.

The percentage increments for nose height were first tabulated in age-sex classes and then subgrouped on the basis of whether in the period between the eighth and fourteenth years the consecutive values for individuals revealed presence (Subgroup A) or absence (Subgroup B) of acceleration. The second and fifth velocity trends in Figure 4 are drawn to age-sex means for the two girls and three boys of Subgroup A; the third and sixth trends to means for the five girls and four boys of Subgroup B.

Considering both sexes together, individual velocity trends for nine of the 14 children (more than 60 per cent) afford no evidence of a circumpuberal acceleration in nose height. The magnitude trends of these children are distributed as follows: concave, four; rectilinear, three; rectilinear-concave, two. Of the five children registering acceleration, only the two girls have similar velocity trends for nose height and stature. It follows that the magnitude trends of these girls are concave-convex-concave. The integral counterparts of the rising velocity trends for nose height on boys are concave-convex (two) and rectilinear-convex (one).

It is concluded that the findings from comparing velocity curves for stature and nose height call in question the generalization that "an adolescent spurt occurs" in "every muscular and skeletal dimension" of the face (27). The discrepancy between such breadth of generalization (5, 20) and the findings of the present study cannot tenably be resolved by recourse to time lag and truncation, i.e., by taking the view that for many of the individuals represented in Figure 4 a circumpuberal acceleration in nose height occurs after the fourteenth year. Nanda (20), from a comparative study of velocity curves for stature and five facial dimensions[4] on 15 children examined repeatedly during childhood and adolescence, finds that where there are circumpuberal velocity increases in facial dimensions their timing is such that "the circumpuberal maximums in height and in the facial dimensions generally occur close to each other."

[4] Not including nose height.

Summary

The problems investigated center on individual growth in skeletal nose height during the postnatal period from age 5 years to age 14 years. Specific objectives relate in part to nose height *per se*, and in part to associations between growth in nose height and growth in other body dimensions.

Highly valid data for nose height are derived from roentgenograms of the face obtained at consecutive annual ages. These are supplemented with longitudinal data for subnasal height, widths of the upper and lower face, head length and width, hip width, and stature. The subjects are North American white children of northwest European ancestry. It is found:

1. For the age span from 5 to 12 years, there are three types of individual magnitude trend for nose height: an ascending curvilinear trend concave downward, an ascending rectilinear trend, and an ascending curvilinear trend concave downward in the first half of the period and convex downward in the last half.

2. Over the seven-year interval between ages 5 and 12 years, the increase in nose height for some children is twice that for other children. This range of variability holds with respect to both absolute and relative increments.

3. There is little, if any, relationship between the growth rate of nose height in early childhood (5 to 7 years) and its growth rate in late childhood (7 to 12 years). The obtained *r*'s, computed from centimeter and percentage gains, do not allow rejection of the null hypothesis.

4. For the childhood period from 5 to 10 years, growth in nose height is not highly related to growth in any of the other variables studied. Obtained associations with all seven traits lie in the zone between low positive and zero.

5. More than one-half of a small series of individual velocity trends for nose height, extending from one to three years beyond the age of the peak of the circumpuberal spurt in stature, give no indication of a circumpuberal spurt in nose height. The impact of this finding on current growth generalization is noted.

References

1. Allen, I. Facial growth in children of five to eight years of age. *Hum. Biol.*, 1948, 20, 109-145.

2. Ashley-Montagu, M. F. The determination of the naso-frontal suture and the nasion in the living. *Radiology*, 1937, 28, 473-476.

3. Baldwin, B. T. Physical growth and school progress. *U. S. Bur. Educ. Bull.*, 1914, No. 10 (Whole No. 581).

4. Bowditch, H. P. Comparative rate of growth in the two sexes. *Boston Med. & Surg. J.*, 1872, 10, 434-435. (See also, Report Brit. Assoc. Adv. Sci., 1880, 50, 140-147.)

5. Davenport, C. B. Post-natal development of the human outer nose. *Proc. Amer. Philosophical Soc.*, 1939, 80, 175-355.

6. Goldstein, M. S. Changes in dimensions and form of the face and head with age. *Amer. J. phys. Anthrop.*, 1936, 22, 37-89.

7. Gray, H., & Ayres, J. G. *Growth in private school children.* Chicago: Univer. Chicago Press, 1931.

8. Herskovits, M. J. The anthropometry of the American Negro. *Columbia Univer. Contr. Anthrop.*, 1930, No. 11.

9. Higley, L. B. Head positioner for scientific radiographic and photographic purposes. *Int. J. Orthodont.*, 1936, 22, 699-704.

10. Holcomb, A. E., & Meredith, H. V. Width of the dental arches at the deciduous canines in white children 4 to 8 years of age. *Growth*, 1956, 20, 159-177.

11. Knott, V. B. Physical measurements of young children. *Univer. Iowa Stud. Child Welf.*, 1941, 18, No. 3.

12. Krogman, W. M. Dental arch form and facial growth pattern in healthy children from prehistoric populations. *J. Amer. dent. Ass.*, 1938, 25, 1278-1289.

13. Meredith, H. V. Growth in head width during the first twelve years of life. *Pediatrics*, 1953, 12, 411-429.

14. Meredith, H. V. Growth in bizygomatic face breadth during childhood. *Growth*, 1954, 18, 111-134.

15. Meredith, H. V. A longitudinal study of change in size and form of the lower limbs on North American white schoolboys. *Growth*, 1955, 19, 89-106.

16. Meredith, H. V. Longitudinal anthropometric data in the study of individual growth. *Ann. New York Acad. Sci.*, 1955, 63, 510-527.

17. Meredith, H. V., & Carl, L. J. Individual growth in hip width: A study covering the age period from 5 to 9 years based upon seriatim data for 55 nonpathologic white children. *Child Develpm.*, 1946, 17, 157-172.

18. Meredith, H. V., & Hopp, W. M. A longitudinal study of dental arch width at the deciduous second molars on children 4 to 8 years of age. *J. dent. Res.*, 1956, 35, 879-889.

19. Meredith, H. V., & Meredith, E. M. Annual increment norms for ten measures of physical growth in children four to eight years of age. *Child Develpm.*, 1950, 21, 141-147.

20. NANDA, R. S. The rates of growth of several facial components measured from serial cephalometric roentgenograms. *Amer. J. Orthodont.*, 1955, 41, 658-673.

21. NEWMAN, K. J., & MEREDITH, H. V. Individual growth in skeletal bigonial diameter during the childhood period from 5 to 11 years of age. *Amer. J. Anat.*, 1956, 99, 157-187.

22. QUENOUILLE, M. H. *Associated measurements.* New York: Academic Press, 1952.

23. SCAMMON, R. E. The first seriatim study of human growth. *Amer. J. phys. Anthrop.*, 1927, 10, 329-336.

24. SHOLL, D. A. Regularities in growth curves, including rhythms and allometry. In E. J. Boell (Ed.), *Dynamics of growth processes.* Princeton: Princeton Univer. Press, 1954. Pp. 224-241.

25. SMYTH, C., & YOUNG, M. Facial growth in children. *Med. Res. Council,* Special Report Series, 1932, No. 171.

26. TANNER, J. M. The assessment of growth and development in children. *Arch. Dis. Child.*, 1952, 27, 10-33.

27. TANNER, J. M. *Growth at adolescence.* Springfield: Charles C. Thomas, 1955.

28. WYLIE, W. L. The relationship between ramus height, dental height, and overbite. *Amer. J. Orthodont.*, 1946, 32, 57-67.

THE INFLUENCE OF VIOLENCE IN THE MASS MEDIA UPON CHILDREN'S ROLE EXPECTATIONS[1]

ALBERTA ENGVALL SIEGEL
Stanford University

With the arrival of television upon the social scene, there has been a renewal of widespread concern that the prevalence of violent themes in the dramatic presentations featured by the mass media may have harmful effects on the audience, especially the child audience. This concern had earlier focused on violent presentations in radio and in movies, and it has been expressed in countless articles and discussions in the popular press and even in the conduct of a legislative investigation of the entertainment industry (10).

Among professional workers, discussions of the possible effects of violence in the mass media have centered on the effects upon children's emotional development—upon anxiety, fears, aggressive motivation, etc. Recent experimental research has not confirmed hypotheses derived from this approach: Maccoby, Levin, and Selya (5) found no significant differences between frustrated and nonfrustrated children in their recall of aggressive movie content, and Siegel (8) found no significant differences in level of aggression or level of anxiety in children after they had seen an aggressive film and after they had seen a comparable nonaggressive one.

Much less attention, however, has been given to the possible influence of violence in the mass media upon what may be called children's *cognitive* development, which includes their conceptual learning of the realities of social life. It is with that possible influence that the present paper is concerned.

Among recent theoretical formulations concerning the variables of child behavior, Baldwin's has emphasized the importance of cognition as well as motivation in the instigation of behavior (1, pp. 112-130). Krech and Crutchfield have also stressed how importantly "cognitive contents" con-

[1] This study was conducted while the writer was the Josephine Glasgow Fellow of the American Association of University Women, to whom sincere gratitude is expressed. The writer is greatly indebted to Ellen Culbertson for participating in many aspects of the conduct of this study, to Professor R. D. Reifsnyder and his students at The Pennsylvania State University for producing the radio dramas, to Harold R. Glenn, Mabel Bresler, and Joan Wood of the Ferguson Township School in Centre County, Pennsylvania, for their generous administrative cooperation in the execution of the study, and to Jean Fuschillo, Joen Fagan, and Irma Shepherd for aid in the collection of data.

tribute to the direction of behavior (3, p. 248). And in classifying personality variables, McClelland includes in his basic triad "the individual's system of schemata," which encompasses "the individual's beliefs, frames of reference, major orientations, role perceptions, ideas, and values" (6, p. 242).

The purpose of the present experiment is to determine whether violence in the mass media influences the nature of that important aspect of personality to which the above cited terms refer. Specifically, the experiment concerns the influence of violence in the mass media on children's role expectations. Role expectations were chosen from the various sorts of cognitive contents enumerated by McClelland (above) not only because they are a sort which might well be expected to be influenced by the mass media but also because they are susceptible of operational definition in research with child subjects.

The concept of role is now a familiar one to behavioral scientists. Linton (4) has defined *status* as position in the social structure, and *role* as the behavior expected from any holder of a given status. Members of any society occupy several statuses—an age status, a sex status, an occupational status, etc.—and culturally assigned to each status is a pattern of expected behavior: a role.

An important function of the socialization process is to train the child to perform the behaviors (roles) associated with his various statuses. The socialized child in nursery school, for example, has learned the role behavior expected of him according to his sex, his age, perhaps his ordinal position in his family, and, to some extent, his socioeconomic status.

However, social life requires more. In addition to being familiar with the behavior requirements associated with one's own statuses, one must also know what behavior is associated with the various other social statuses, in order that one may interact with persons occupying these. To function effectively, the nursery school child must know what behaviors are to be expected from children of the opposite sex, from adults in the statuses of teachers, parents, "observers," "student teachers," etc. To a large extent, the adequacy of a child's social behavior will depend on the appropriateness of his "cognitive contents" concerning the role requirements of various social positions.

In accounting for children's behavior, we commonly refer to some postulated motive as explanatory. For example, an aggressive child may be said to have strong aggressive motivation. While the relevance and usefulness of such concepts cannot be denied, it is important to consider as well the child's cognitive learning in our attempts at explanation. If a child has learned to think that all spiders are dangerous, we need not attribute his destructive behavior toward a spider to any "aggressive motive." Moreover, aggression between children may sometimes be accounted for in terms of cognitive learning: either the child's concept of his own role may call for aggression, or perhaps his concept of the other child's role

leads him to expect aggression from the other and therefore arouses his own aggressive behavior in defensive offense.

This discussion is not to deny the importance of motivation in the instigation of behavior, but rather to call attention to the existence of other sources of instigation. These sources may well interact and modify each other, but it is important that conceptually we distinguish among them.

The term "role expectation" (7, p. 226) denotes a person's knowledge of what rights and obligations, actions and qualities, are appropriate to persons of a given social status. In the socialization process, role expectations are acquired not only through intentional instruction but also through incidental learning. Cultural agents may intentionally instruct a child concerning behaviors appropriate to his own statuses (age, sex, grade in school, socioeconomic level, religious affiliation, etc.) and occasionally concerning behaviors appropriate to the statuses of others. In addition, role expectations will be learned incidentally from a variety of sources. Incidental learning may be especially important in learning role behaviors appropriate to statuses held by others or to statuses which the child may himself hold at some future date. The sources of incidental learning about roles include not only living persons occupying various statuses (role models), but also literature, myths, jokes, stereotypes, etc., as well as explicit statements of social norms such as the Boy Scout Code (6, pp. 314-315).

To date, our information concerning the learning of role expectations comes from observation, recall, and other informal and uncontrolled sources. "Controlled experimentation on the learning of roles has not yet begun" (7, p. 227). The present study is an attempt at controlled experimentation in this field, and concerns one, possibly very potent, source for incidental learning of role expectations: dramatic presentations in the mass media.

The hypothesis under test is that role expectations held by children are modified by the content of dramatic presentations in the mass media. That is, the role behavior exhibited by various characters in dramatic presentations may be incorporated in the child's role expectations for actual persons occupying the statuses represented by the characters.

Many informal observations support this hypothesis. When Eastern children, especially those from urban areas, vacation in the West, they often expect ranch hands to behave like the cowboys they have seen so often in Westerns on TV and at the movies. Their fearful expectations of Indians reflect their faithful acceptance of the portrayals of Indians in these same sources. Closer to home, children commonly fear policemen, and this may in part be a result of their acceptance of the stereotyped portrayal of policemen in the mass media, wherein they appear as hostile and aggressive.

The general hypothesis will require many tests of its validity and generality. The extent to which children accept as socially accurate the role portrayals they see and hear in the mass media doubtless depends on such variables as: the extent and quality of their personal experiences with the role being portrayed; the consonance between the mass media portrayal

and other portrayals; the realism or irrealism of the dramatic form used in the mass media presentations; the age, maturity, experience, etc. of the child audience.

The present study provides one specific test of the hypothesis by using a highly "realistic" form of dramatic presentation, using as subjects relatively immature and unsophisticated children, and focusing on a social status which is familiar to the child subjects but for which they have typically not developed highly structured expectations.

Experimental Method

Experimental confirmation of the hypothesis would be provided by a demonstration that different expectations of the same role will follow exposure to different dramatic presentations of that role in the mass media. Since this result is most likely in the case of role expectations which are not already definite or firmly structured, it was necessary to focus this experiment on a social role concerning which the subjects had no strong preconceptions. The subjects of the experiment were to be second graders, and therefore preliminary individual interviews were conducted with a small group of second graders (not members of the subject group) to determine their role expectations for a variety of social statuses held by men. (The pronounced sex differences in aggression in our culture led to the requirement that the roles be ones typically played by men. It was felt that any attempt to induce the subjects to include aggression in their role expectations for some women's status would have conflicted with the children's prior learning about sex-appropriate behavior.) From these informal discussions, the role of taxi-driver emerged as one with which the second graders were familiar, but concerning which they had no firm expectations or definite stereotype. Accordingly, the experiment used role expectations of taxi-drivers in the test of the hypothesis.

In brief, the experiment was conducted by exposing two equivalent groups of children to different series of dramatic presentations of the role of taxi-driver, on a medium of mass communication. Following these exposures, the children's role expectations of taxi-drivers were assessed in a different context. The hypothesis was that these expectations would differ in accordance with the contents of the dramatic presentations. The procedure is detailed below.

Subjects

All participants in the study were second graders in a nine-grade elementary school which is located in a small village and draws its students from that village and the surrounding rural countryside. The parents of the children were farmers, skilled and unskilled workmen, operators of small businesses, and institution employees.

In the school, the second graders were enrolled in two classrooms. These classes constituted the experimental (E) and control (C) groups of the present study. The acceptance of existing groupings for use in the experiment was justified on the ground that school administrators had assigned children to the two classes on a random basis; it permitted introducing the research procedures as part of the regular school routine. The descriptive data given below suggest that the E and C groups do not differ systematically, but rather are from a common population, as would be expected from the manner in which they were initially composed.

In the E classroom, 29 children were subjects of the study, having qualified by being in attendance for the dramatic presentations and the subsequent testing. In the C classroom, in which fewer second graders were enrolled, 15 children were subjects of the study.

The median age of the E subjects was 7-7, while that of the C subjects was 7-6. In both groups, girls were numerically preponderant: the E group included 18 girls and 11 boys, the C group, 9 girls and 6 boys.

The school records included intelligence test scores for 37 of the subjects. These scores were based on testing conducted nine months prior to the experiment, late in the subjects' first-grade year in school. Thurstone and Thurstone's Primary Mental Abilities test for ages 5 to 7 (9) was employed. This test yields IQ equivalents, normalized standard scores with mean 100 and standard deviation 16. For the E group, 23 first-grade IQs were available. These ranged from 93 to 118 with a mean of 107.2. For the C group, the 14 IQs which were available ranged from 93 to 119, with a mean of 107.6.

Both the ages and the IQs of the subjects suggest that the E and C groups were drawn from the same population. Acceptance of this proposition, which is based on statements of school administrators responsible for class composition as well as on the above data, permits us to attribute any observed post-experimental differences between the E and C groups to the experimental condition rather than to possible prior group differences.

Dramatic Presentations

Because exact control was required over the contents of the dramatic presentations on a medium of mass communication, those used in this study were written and produced especially for it. For reasons of feasibility in both production and presentation, radio dramas were used rather than movie or TV dramas.

Three radio dramas were written. Each was presented as one of a continuing series, "Taxi Driver." Each production was presented at the same time of day, and each used the same theme music, format, announcer, and introductory and concluding statements by the announcer. Thus, the productions had the usual characteristics of a series of radio productions. A taxi-driver was a central character in each of the three dramas, but the identity of the taxi-driver was different in each, for the purpose of the

study was to focus attention on the taxi-driver role itself, not on the possibly idiosyncratic characteristics of any particular person occupying that status. Each of the three productions employed devices to clarify the occupational characteristics of the taxi-driver role: the announcer discussed what a taxi-driver does, the character himself carried on brief conversations with his customers on this topic, and the story line incorporated the occupational characteristics of a taxi-driver. These devices were employed to insure that all of the listeners shared some common conception of what a taxi-driver does ocupationally.

Each of the three dramas was written in two versions, an E and a C version. That initial part of each drama which set the scene, introduced the characters, and propounded the dramatic conflict was common to both versions. The two versions differed only in the *outcome* of the story: in the E version, the dramatic conflict was resolved by aggression and physical violence on the part of the taxi-driver, whereas in the C version another (nonaggressive) solution occurred.

The three dramas, each in two versions, were enacted and tape-recorded by students in an advanced class in radio production. For each drama, the participants in both the E and C versions were the same actors; in fact, the initial portions of the two versions, being common, were enacted but once and then this enactment was transferred to both the E and C tape recordings.

Content

Each of the six productions began with the theme music. The announcer's voice then broke in to identify the program, "Taxi Driver." When the music faded, the announcer made this opening statement:

> Here is another story about a man who drives a taxi-cab. Every day, in every part of the country, many men make their living by driving taxi-cabs which carry other people. Their job is to drive the taxi. Taxi-drivers take people to work. They take passengers to meet trains. Taxi-drivers drive women downtown to go shopping. Sometimes they deliver children to their schools. A taxi-driver may drive a patient to see a doctor or to go to a hospital. A taxicab-driver has many adventures in his work. This program tells about adventures in the everyday lives of taxi-drivers.

After this standard statement, more of the theme music was played, and then the announcer introduced the day's episode. His introduction was followed directly by the enactment of the drama, which required about 10 minutes. At its conclusion, the theme music was heard again, and then the announcer returned with:

> And that ends today's story of everyday adventures in the life of a taxi-driver. Listen in tomorrow at the same time for another thrilling story from the life of a taxi-driver!

After this statment, the theme music returned for a final time, completing the production.

Limitations of space preclude summarizing each of the three dramas here. Given below is a synopsis of the third drama, which will serve to exemplify the nature of all three. Its initial portion, common to both E and C versions, is given first, followed by the two alternative outcomes.

Drama III

Bill Miller is taking his woman taxi passenger home from Christmas shopping, and they discuss the problems of driving in icy weather. When she comments on the difficulty of his job, Miller launches into a lengthy boast about his perfect driving record, unbroken in years of fulltime taxi driving. The talk turns to another topic, but is interrupted by a loud crash, accompanied by shrieking brakes and screams of fright: the taxi and another car have collided. After determining that his passenger is not injured, Miller steps out to confront the driver of the other car.

E outcome: The two drivers blame each other for the accident. Miller becomes incensed at this accusation, citing his own perfect driving record. The other driver persists in blaming Miller, and suggests that he may petition for Miller's license to be revoked. At that, Miller threatens physical violence. When the other driver fails to yield in the face of Miller's abuse and threats, Miller makes good his threat to "beat you bloody." The passenger is shocked at the fight; when Miller returns to the cab, she expresses alarm that the other driver is lying motionless on the ground, hurt. Miller's only reply is, "I hope he *is* hurt." He turns his attention to the taxi. The passenger decides to walk to her destination, and Miller takes his cab to headquarters for repairs. There he boasts about having beaten up the man whose car hit his taxi.

C outcome: The two drivers begin by blaming each other for the accident, but Miller is mollified when he learns that the other driver is from California and is therefore unfamiliar with the techniques of driving over icy roads. When they turn their attention to their cars and to obtaining a report for insurance purposes, the passenger is consulted. In the course of the conversation, she recognizes the driver of the other car as her brother-in-law, with whom she had been out of touch for 20 years. Thoughts turn from the crash to the forthcoming reunion between her and her sister. Miller comments that he doesn't mind so much that his perfect driving record was broken, now that the crash has served to reunite a family.

Administration

The three dramas were presented to the children on three successive days (Monday, Tuesday, and Wednesday) at the same time-of-day. The purpose of presenting each drama at the same time of day was to preserve the illusion that the dramas were authentic radio broadcasts. The "radio" was a console which housed the tape player, and which was constructed to have the external features of a radio. It was wheeled in and out of the room for each presentation. The children were told that they were being given the opportunity to hear the programs so that the person who brought the radio to the room each day could determine whether these were the sort of programs enjoyed by second grade children.

Informal observation during the radio presentations indicated that the children enjoyed the dramas, typically became "involved" in them, and in no way questioned their authenticity.

Measure of Role Expectations

To assess the concept which each child held of the role of taxi-driver after hearing the radio dramas, a test was administered individually to each child. Testing was conducted on the two days following the final radio presentation, on Thursday and Friday of the same school week. The test, described below, was a completion test, orally administered on an individual basis, in which the child was asked to give completions to selected newspaper stories.

Rationale

The role expectations test was designed to elicit a child's *social* or *reality* conception of the role of taxi-driver, with special reference to his conception of the taxi-driver's resort to violence in his interpersonal relations in conflict situations. In certain structural features, the role expectations test bears a superficial similarity to projective tests of the story completion type (cf., 2). The nature and rationale of the role expectations test may be clarified by a discussion which contrasts it with tests of the projective type.

In a projective test, a primary aim is to tap the child's wishes, impulses, fantasies, motives, etc.; it is therefore necessary to minimize "contamination" of his responses by his conceptions of social realities, of what behavior is socially appropriate or standard. To accomplish this, certain techniques are typically used:

1. The test is introduced as a measure of imagination or creativity.

2. The instructions inform the subject that there are no "right" or "wrong" test responses.

3. In the stimulus material of the test, social reality features are minimized. In a picture test, for example, the status identities of the various characters may be unclear or ambiguous: the pictures are so constructed that the subject is not compelled by the stimulus material to identify any character as being of a particular age, sex, or other status position. Alternatively, social reality features may be avoided, as in some picture tests, by the use of animals rather than humans as story characters.

Each of these techniques is designed to minimize the social compulsion inherent in the stimulus material and therefore to maximize the subject's projection of his own impulses and wishes. The test material contains a minimum of features which would impose structure on the subject's response, and would therefore interfere with, or color, the operations of projection. The test instructions are formulated to serve the same end.

In the role expectations test used in this study, on the other hand, the focus is on the products of the child's social learning, and specifically on his understanding of what behavior is role-appropriate to a particular social position. Therefore, the technique aims to *minimize* the child's projection of his own wishes and needs onto his response. For this test, projected material can only be conceived as "contaminating." To minimize projec-

tion, several procedures are used which may be contrasted with those mentioned above:

1. The role expectations test is introduced as a measure of achievement, of how much the child knows.

2. The child is given to understand that there are "right" and "wrong" answers, and thus his own responses will be judged in relation to the correct ones.

3. Social reality aspects are *maximized* in the test materials. The stories to be completed are newspaper stories, and they are presented as being from a current edition of the local newspaper. The characters in the critical stories are identified according to their sex and other status positions, and the events of the stories are cast in a specific, concrete, and familiar local context.

Administration

The test was administered orally to each child in individual sessions. Testing was conducted in small private rooms in the school building. Two women with whom the children had no prior acquaintance administered the test. They followed a common procedure, having conferred in the development of standard techniques prior to the experiment.

Every effort was made to disassociate the testing from the radio dramas. The person who had brought the "radio" to the classrooms each day was not one of the testers, and was never seen at the school in either tester's company. Classroom teachers made no mention that the radio dramas and the testing had any relation to each other. The two testers did not appear at the school at any time before or during the presentation of the radio dramas. They made no reference to the dramas in their conversations with the children. If a child mentioned the radio dramas to the tester, she responded as if she had not heard of them before.

An assistant brought each child to the door of the testing room and introduced him to the tester. This procedure prevented the tester's knowing whether any child was from the E or C group.

When the child was seated, the tester said, "Jimmy, tell me what a newspaper is. What is a newspaper? What does it do?" The purpose of this opening inquiry was to determine whether the child understood that a newspaper reports events of the "real world," as distinct from presenting fiction. The child's response to this question was recorded verbatim, after which the tester entered a rating of either 0 or 1 in the child's protocol, indicating her judgment as to whether or not the child comprehended that a newspaper serves to report social reality.

In responses which received a rating of 1, it was common for the child to cite specific sorts of news which commonly appear in newspapers, and even to mention particular events which had been reported. One child mentioned that his sister's name was in the newspaper when she broke her arm. Another mentioned a news story reporting a fire in a neighbor's barn.

Many mentioned that automobile accidents are reported in newspapers. These responses are typical: "You get news in it. Like a wreck or something—you see it in the paper. Pictures too. You read in there what happened." "Tells what's coming up. Tells the weather. Wrecks and things." "It comes in the mail. You read it. It tells about things that happen: people get a baby, someone dies, someone gets hurt. My brother had his finger cut off and that was in the paper."

A child received a rating of o if he was unable to answer the inquiry, if his response merely described the physical characteristics of a newspaper, or if his response otherwise gave no indication that he understood a newspaper's reality reporting function. For example, one child said only, "You read it." Another said, "A paper, about fires and houses falling down," and another, "A thing that you read about things, like a wreck."

After the child had answered the initial question, the tester described the function of a newspaper, making this description sound like a confirmation, restatement, or extension of what the child had said. "Um-hum, a newspaper tells us what happens in the world, doesn't it? It reports the news: the things that happen to people. It tells about the activities of people in our community and in other places." This statement was followed by a pause to permit the child to make further remarks about newspapers.

Then the test was introduced.

> Now, Jimmy, today I want to give you a little test. I want to find out how much you know about what is going on in the world. The way I'll do that is to find out about what you know about the news that is reported in the newspaper.
>
> See these newspapers? They are copies of the *Centre Daily Times,* the newspaper for our area. I have chosen some news reports from these newspapers. I want to find out what you know about these reports. Here is what we will do. I will read to you the first part of a news report. Then I will stop, and you tell me the rest. You tell me how the situation came out. Now sometimes you may not know exactly, and then I want you to guess. If you do not know too much about the news story I am reading to you, then tell me how you would expect such a situation to work out.

This statement was made slowly to permit the child to respond as he chose, and statements which seemed unclear or confusing to the child were rephrased or repeated.

The tester had copies of two current issues of the local newspaper. On the inside pages of these, six "bogus" news stories, written for the purpose of this study, had been pasted. The tester was careful to hide these stories from the child as she turned the pages of the newspaper and read the stories. This procedure is not unusual in an achievement test, and neither tester was aware that any child suspected the authenticity of the news stories.

The six stories are given below in the order in which they were read. Subjects' completions to each were recorded verbatim.

The first two stories, concerned with events with which the children were familiar, were intended to give each child initial "success" experiences and also to confirm to him the reality-oriented nature of the test. These stories set the stage for the remaining four, which were entirely fictional and to which the child could not "know" the answers. In order to elicit acceptable responses to the first two stories, testers gave some prompting, if necessary, but their techniques were designed so that that child could think he had reached a correct answer independently. Responses to stories 1 and 2 were reinforced by the tester's enthusiasm and comments like "That's right," "Yes, it was cold, wasn't it?"

The testers gave no prompting or specific reinforcement for completions to stories 3 through 6. The initial instructions had encouraged the child to guess on stories for which he did not know the ending, and if necessary these instructions were repeated after story 3 had been read. To encourage the subjects, the testers used such techniques as these: nodding, saying "um-hmm," expectant pauses, questioning looks, holding pencil poised, etc.

Content

The six news stories used in this study were selected from a larger group which had been pretested with the pilot study group of children. Selection was based on the apparent comprehensibility and interest of the stories.

Story 1

There will be no school next Monday. All public schools will be closed, and school children will have a holiday. The reason that next Monday is a school holiday is . . . . What is the reason?

The correct answer was that Monday was Lincoln's birthday. To aid the research, teachers in both classrooms had conducted discussions of that forthcoming holiday. However, some children suggested that the holiday was to celebrate Valentine's Day, Washington's birthday, Eisenhower's birthday, or Roosevelt's birthday, and these responses were accepted and praised.

Story 2

Winter weather has been visiting Centre County for the last few days. The roads have been slippery and drivers have found driving difficult. The sidewalks have been slippery too. In Pine Grove Mills (the name of the town in which the school is located), the weather has been . . . . Now tell the news about the weather here.

These are typical responses: "Bad. Rainy. Snow." "Snowy and icy." "Kinda bad. We had some wrecks. People have been falling down and hurting themselves."

Story 3

Bill Smith was walking down Allen Street today, and he saw some money lying on the sidewalk. He walked over to pick it up. Just then Howard Brown stopped him. Brown said that he had lost the money earlier

in the day, and that he had been looking for it for a long time. Brown said that he was sure the money was his. Smith asked Brown to prove that this was the money he had lost. Brown could not prove it. So then Smith said that he was going to keep the money, because he had seen it first. . . . . Now tell me what happened then.

Story 4

When John Henry got home from work today, he found a TV man there installing the new television set which the Henrys had ordered. As soon as the set was installed, the whole Henry family gathered around to watch television. They were excited about having their own TV set at last.

But neither Mr. Henry nor his wife could get the set to work. No matter what they tried, they couldn't get the TV set to work at all. Finally Mr. Henry called the TV company and told them to send the installation man back to the Henry house. The man came, and he couldn't make the set work either. When that happened, Mr. Henry told the TV man that he should take back the TV set and give them a good one instead. The man said that he couldn't do that. He said it was too late for Mr. Henry to get a different set. He also would not give back the money the Henrys had paid. The TV man told Mr. Henry, "You have already paid for that set and now you're stuck with it, even if it is no good!" . . . . What happened then?

Story 5

Today was a busy day for taxi-drivers, for cold weather always brings extra business for taxi-cabs. But John O'Neal didn't earn as much money this morning as he might have. O'Neal, a taxi-driver for the local taxi company, was called to pick up a passenger at his home. O'Neal arrived at the address with his taxi at nine o'clock, and honked his horn to let the customer know that the taxi was ready.

The passenger did not come out, and O'Neal waited for him in his taxi for 15 minutes. Finally he went to the door of the house and knocked, in order to see what was keeping the passenger. The man came to the door and said he would be ready to go in just a minute. O'Neal went back to his taxi and waited there in the cold for another 20 minutes, which is a long time. Of course, he was losing money all this time, because he wasn't doing any business. Finally he went to the door and knocked again. This time the man came to the door and told the taxi-driver that he had decided to stay home and therefore he would not need the taxi. . . . . What happened then?

Story 6

The cold and slippery weather brought bad luck today to Fred Horner, a taxi-cab driver. Up until today, Horner has had a perfect driving record: he never had an accident while driving his taxi-cab. But today there was so much ice on the streets that it was very hard to drive. So there was a crash between Horner's taxi-cab and another car. That crash broke Horner's perfect driving record. Both cars were damaged in the accident, but no one was hurt. After the crash, the taxi-driver got out of his taxi and . . . . What happened then?

The reader will recognize that story 6 presents the same conflict as was presented in the initial portion of drama III. In this, story 6 differs from story 5, which is a "new" story—its plot does not duplicate or resemble the plots of any of the three radio dramas.

Scoring

Two methods of scoring were devised to permit two approaches to testing the research hypothesis. The first mode involved a comparison of the total amount of aggression included in a child's completions to stories 5 and 6 (the stories in which taxi-drivers figure) with the total amount included in his completions to stories 3 and 4. The second mode focused on the amount of aggression attributed to the taxi-driver in each child's completions to stories 5 and 6.

Difference in aggression. The first mode of scoring took into account children's individual differences in the amount of aggression they attribute to others. As would be expected, some children included aggressive incidents in every completion they gave, others included no aggression in any completion, and others ranged between these extremes. To eliminate the influence of such individual differences on scores used to test the hypothesis, a difference judgment was obtained.

Each child's completions to stories 3, 4, 5, and 6 were assembled. A judge read the four together, and compared the amount of aggression shown in completions 3 and 4 with the amount shown in completions 5 and 6. The latter stories concern taxi-drivers; the former do not. Any sort of aggression, from any source, was considered in this evaluation. The judge then assigned the child's protocol to one of two categories: *plus* or *zero-minus*. The *plus* category was for protocols whose completions to stories 5 and 6 contained more aggression than the completions to 3 and 4. The *zero-minus* category was for protocols whose completions to the two sets did not differ in aggression and for protocols in which completions 3 and 4 showed more aggression than completions 5 and 6.

Two judges made these categorizations independently without knowing which children were from the E and which from the C group. Their two sets of judgments showed 92.5 per cent agreement (total number of agreements / total number of judgments).

Taxi-driver aggression. Completions to stories 5 and 6 were also evaluated by another method. They were categorized according to the amount of aggression attributed to the taxi-driver.

The completions to the two stories were evaluated separately. For the judging, the E and C protocols were intermixed, and the judge had no knowledge of the group from which any protocol came. Judges were instructed to assess only the aggression attributed to the taxi-driver, not the total aggression attributed to all sources. Each completion was assigned to one of three categories, according to the amount of aggression attributed to the taxi-driver. A completion was assigned to the *high* category in cases where the taxi-driver fought, inflicted injury or pain, reported other persons to the police, got mad, yelled, was angry. A completion was assigned to the *intermediate* category if the taxi-driver showed some aggression (complained, accused, argued, etc.), or if the response implied that others

may fear a taxi-driver even though no aggression is ascribed to him, or if it is implied that opposing a taxi-driver is dangerous and may well bring trouble. A completion was assigned to the *none* category if no aggression was shown by the taxi-driver in it.

Each of the 44 protocols was scored independently by two judges to permit an assessment of the reliability of the scoring system. For completions to story 5, the judges were in complete agreement in categorizing 42 of the protocols, and differed by but one category in the other two cases. For completions to story 6, again the judges differed in only two cases, and then again by but one point. Thus, for each story, the judges showed better than 95 per cent agreement in their categorization of the completions.

Summary of Experimental Controls

Certain controls were included in the procedure to clarify the meaning of the outcome of the research. They have been mentioned in various contexts above, and are summarized here:

1. The E and C dramas were equivalent in format, acting personnel, length, initial portion, and in the manner in which they were introduced to the children. They differed in outcomes, in the attribution or nonattribution of aggression to the taxi-driver.

2. All evidence supported the assumption that the E and C subjects were from a common population, so that post-experimental differences need not be attributed to prior group differences but rather may be attributed to the treatment condition.

3. Every effort was made to disassociate the radio presentations from the subsequent testing.

4. Testers were not able systematically to identify the group (E or C) from which any child came when he was brought to the testing room.

5. Judges who scored the protocols had no extraneous information concerning the group (E or C) from which any protocol came.

Results

The first step in the treatment of the data was to eliminate from consideration those children who gave evidence of not understanding the reality-reporting function of a newspaper, because these children's completions to the newspaper stories in the test could not be interpreted as reflections of their understanding of social reality. Children whose responses to the question "What is a newspaper?" were given a rating of 0 by their testers were dropped from the analysis. Four subjects received this rating. All four were boys, three from the C group and one from the E group. After they had been dropped from the subject group, the E group consisted of 10 boys and 18 girls while the C group consisted of 3 boys and 9 girls. Except in the ratio of boys to girls, the elimination of these subjects

did not render the E and C groups less equivalent. After the dropping of cases, the mean IQ in the E group was 108.5, in the C group, 108.4; the mean age in the E group was 7-7, in the C group, 7-6.

The mode of collecting the data permitted two approaches to testing the research hypothesis that the radio presentations would influence the children's role expectations of taxi-drivers, causing the E children to be more likely to attribute aggression to the taxi-driver role.

Difference in Aggression

The first test of the hypothesis utilized the difference judgments by which each child's completions to stories 5 and 6 (the taxi-driver news stories) were compared to his completions to stories 3 and 4. The prediction was that the proportion of subjects assigned to the two categories (*plus* and *zero-minus*) would differ in the E and C groups, with the E group being positively displaced relative to the C group. That is, the prediction was that, compared to the C group, the E children would have a higher incidence of aggression in completions 5 and 6 than in completions 3 and 4.

(If we could assume that stories 3 and 4 are *a priori* equivalent to stories 5 and 6 in the amount of aggression likely to be contained in children's completions to them, then we could simply predict that *plus* ratings would be more frequent than *zero-minus* ratings among the E children, because of the experimental treatment. Such an analysis would use each child as his own control and would not require the use of any separate control group. Since there is no good information to support the assumption that the two sets of stories are *a priori* equivalent, a comparable test is made by determining whether the E and C groups differ in their proportions of *plus* and *zero-minus* ratings.)

The data for this test are given in Table 1, which shows that the groups diverge in the predicted way. A chi-square test of these data shows the difference to be significant: with Yates correction for continuity, $\chi^2 = 3.29$, $p < .05$, one-tailed test. Only one-sixth of the C children had more aggression in their completion to the taxi-driver news stories than in

TABLE 1

DIFFERENCE IN AGGRESSION BETWEEN COMPLETIONS TO TAXI-DRIVER NEWS STORIES AND TO NON-TAXI-DRIVER NEWS STORIES BY E AND C SUBJECTS

	C Subjects	*E Subjects*	*Total*
Difference in Aggression			
Plus	2	15	17
Zero-Minus	10	13	23
Total	12	28	40

their completions to the others, while the majority of the E children had more. This is evidence for the effectiveness of the radio dramas in shaping the children's expectations of taxi-drivers.

Taxi-Driver Aggression

The second mode of analysis provides a more specific test of the research hypothesis, taking into account only that aggression attributed to taxi-drivers by the E and C children in their completions. This analysis is accomplished by a direct comparison of the E and C children's attribution of aggression to taxi-drivers in completions to stories 5 and 6. The data are the categorizations in terms of taxi-driver aggression: children's completions to these two newspaper stories were categorized as *high, intermediate,* or *none,* in terms of the amount of aggression attributed to the taxi-driver. The prediction was that the E children's completions would more frequently attribute aggression to taxi-drivers than would the C children's. The data for the test of this prediction are given in Tables 2 and 3.

TABLE 2

AGGRESSION ATTRIBUTED TO TAXI-DRIVER IN COMPLETIONS TO NEWSPAPER STORY 5

	C Subjects	*E Subjects*	*Total*
Aggression of Taxi-Driver			
High	4	8	12
Intermediate	3	3	6
None	5	17	22
Total	12	28	40

Table 2 gives the data for newspaper story 5. In amount of aggression attributed to the taxi-driver in completions to this story, the two groups show no difference.

Table 3 gives the data for story 6. In amount of aggression attributed to the taxi-driver in completions to this story, the two groups differ significantly in the predicted direction: $\chi^2 = 8.40$, $p < .02$. (The small expected frequencies in some cells in Table 3 suggest that a supplementary test of these data be made. A more conservative analysis was performed by combining the "high" and "intermediate" cells. For the resulting 2×2 table, with Yates' correction for continuity, $\chi^2 = 2.81$, $p < .05$, one tailed test.) About half of the E children attributed *high* aggression to the taxi-driver in their completions to story 6, while no C child attributed *high* aggression to the taxi-driver and in fact two-thirds of them attributed *none* to him.

TABLE 3

AGGRESSION ATTRIBUTED TO TAXI-DRIVER IN COMPLETIONS TO NEWSPAPER STORY 6

	C Subjects	*E Subjects*	*Total*
Aggression of Taxi-Driver			
High	0	13	13
Intermediate	4	6	10
None	8	9	17
Total	12	28	40

These data suggest that the main source of the difference in aggression attributed to taxi-drivers by E and C children is the difference in their completions to story 6. This is the story which paraphrased the initial part of radio drama III. Inspection of the contents of the completions to story 6 reveals many instances of literal carry-over from the contents of the radio drama to the contents of the completions. Of the 40 completions, 26 referred to the sex of the driver of the other car, and in every one the driver was identified as male, as was the case in drama III. Eight completions mentioned a passenger in the taxi, and each one identified her as female, as was the case in drama III. Two children mentioned the color of the car with which the taxi collided, and both identified it as blue, as it was in the radio drama. Many children said that the taxi-driver hit (fought with) the driver of the other car, of course, but only two mentioned any specific number of hits, and both gave that number as three, which it was in the radio drama. Six children's completions constituted more or less faithful recountals of the contents of the radio drama, either E or C, whichever they had heard. This one is perhaps the most complete and literal: "He checked his car. He hit the man three times. The lady in the car wanted to pick up the man, but he said 'No, he broke my perfect record.' The lady said she would walk to the store. He said 'Call me if you need help.' Then he drove to the station and reported 'A man came along and broke my perfect record.'" Four children, including two who were in the group just mentioned, made some overt reference to the radio dramas in their responses to story 6.

DISCUSSION

Interpretation of the Results

The interpretation of the results reported above rests on the assumption that the newspaper test has face validity as a measure of children's reality expectations. As the discussion of this test in an earlier section emphasizes,

all of the procedures of the testing were designed to establish this validity. Aimed toward the same end was the procedure of dropping from the analysis those children whose responses to the question "What is a newspaper?" did not indicate that they understood the reality-reporting function of a newspaper.

This is not to say that seven-year-olds distinguish between reality and fantasy with the same clarity as do adults, but rather it is to say that insofar as they do make this distinction they understand that newspaper contents are in the reality sphere. In point of fact, the hypothesis of this study is cogent precisely because the difference between reality and fantasy is so blurred and unclear for young children. The incorporation of many of the incidental features of drama III in many children's completions to story 6 illustrates this naivete.

Two modes of scoring the completions were used, and the statistical analysis of both sets of scores provides confirmation for the hypothesis that the radio presentations would influence the children's role expectations of taxi-drivers, causing the children who heard the dramas with endings featuring taxi-driver violence to be more likely to attribute aggression to the taxi-driver role. The first analysis demonstrated that the likelihood that more aggression would be included in the completions to taxi-driver news stories than in completions to other news stories was greater among the E than among the C children. The second analysis demonstrated that the E children attributed more aggression to the taxi-driver in their completions to one of the taxi-driver newspaper stories than did the C children.

Two interpretations may be attached to the finding that the E and C children did not differ in the amount of violence they attributed to the taxi-driver in their completions to story 5, but did differ, in the expected direction, in their completions to story 6.

Newspaper story 6 was the one which replicated the initial portion of drama III. Newspaper story 5 also concerned a frustration which came to a taxi-driver in the course of his work, but not a frustration which had been depicted in any of the radio dramas.

First, it may be that the children's learning from the three dramas was specific rather than general. That is, perhaps they learned only what is to be expected from taxi -drivers in the specific frustration situations depicted in the three dramas, rather than generalizing from this some inclusive formula like "taxi-drivers respond to frustration with physical aggression." Only when the reality situation duplicated the situation presented in the radio dramas were their expectations based on what had been presented in the dramas; hence, the group differences in response to story 6 but not to story 5.

Second, and alternatively, it may be that the children's learning from the three dramas *was* the generalization that taxi-drivers respond to frustration with aggression, but that the situation presented in newspaper story 5 was not perceived by them to represent frustration to the taxi-driver.

The frustration resulting from having to wait long and fruitlessly for a passenger may seem less severe than the frustrations depicted in the radio dramas, and therefore the formula "taxi-drivers respond to frustration with aggression" was not invoked as relevant.

The first interpretation could be tested by an experiment which gave the children greater opportunity to generalize from the radio dramas. One method would be to include many more radio dramas in the experiment, say 15 or 20, each depicting a different taxi-driver responding with physical aggression to a different frustration.

The second interpretation could be tested in an experiment in which other newspaper stories were used, stories in which the frustration to the taxi-driver is more evident to the children.

The present study, however, provides no specific evidence supporting either one or the other of these alternative interpretations.

Generality of the Hypothesis

The hypothesis that audiences generalize to social life from the contents of the mass media is a familiar one. It has long been recognized that the mass media have powerful "educative" capacities. Many have suggested, for example, that the stereotyped portrayals of Negroes in movie and television drama may serve to perpetuate and reinforce the American myth of ethnic differences, of "racial" superiority and inferiority. Those concerned with international understanding have frequently expressed concern over the representation of American life contained in Hollywood films exported abroad. A suspicion that dramatic portrayals in the movies are naively accepted as representations of reality underlies the fear that citizens of other nations may build a distorted picture of American life from their moviegoing experiences. It was this fear which the United States Ambassador to Italy, Clare Boothe Luce, invoked when she objected to the showing of the film "Blackboard Jungle" at the Venice Film Festival in 1955.

Similar reasoning underlies the argument that American children may acquire from the mass media a distorted conception of the realities of American life, particularly of the uses of violence in social life. This argument is also supported by the notion that children are less sophisticated than adults in distinguishing reality from fantasy.

We lack information concerning the conditions under which the carry-over occurs. Confirmation of the hypothesis of this study in one research context is not sufficient. The hypothesis requires many independent tests of its validity and generality. We need research directed to questions like those below.

Does the carry-over occur for roles with which children have had considerable personal experience as well as for those with which they have had little or none? For example, is the stereotyped portrayal of the "old maid schoolteacher" accepted as readily and unreservedly as the portrayal

of cowboys? Children have firsthand social experience with teachers, but usually not with cowboys.

How potent is a role portrayal in the mass media which runs counter to the role portrayals in other sources of role learning (jokes, stereotypes, literature, social experience, etc.)?

Are role expectations learned from unrealistic and highly fantasied dramas (for example, animated cartoons with animal characters, or stories of fairies, witches, pixies) as well as from realistic ones? We might guess that dramatic presentations in the "documentary" style have greater carry-over than do dramatizations of low realism, such as musical comedies, fairy stories, animated cartoons, etc.

How does the child's acceptance of mass media portrayals depend on his maturity? Young children, lacking experience and not always able clearly to distinguish reality from fantasy, may be more uncritical and unreserved in their acceptance of mass media portrayals than are more mature children.

A full evaluation of the influence of violence in the mass media upon children's reality expectations will require results from experimental approaches to answering the above questions.

Summary

The hypothesis of the present study is that children's reality expectations are influenced by the contents of dramatic presentations in the mass media. More specifically, the hypothesis is that the aggression and violence they include in their role expectations of occupants of certain statuses is influenced by the attribution of aggression and violence to occupants of those statuses in dramatic characterizations in the mass media. Whereas discussion among professional workers concerning the influence of violence in the mass media upon children has centered on possible emotional effects, this hypothesis is addressed to possible cognitive effects.

To test the hypothesis, a series of radio dramas was presented to the children, after which their role expectations of the role held by the central character in the series were assessed by means of a newspaper test.

Subjects were second graders in two classrooms, E and C. The radio dramas were each enacted in two versions; the E children heard those in which the central character resolved the dramatic conflict by personal physical aggression, while the C children heard those in which another solution occurred to the same conflict. The E and C dramas shared a common opening, were enacted by the same personnel, and produced under the same format. They differed only in outcomes. On the days which followed the conclusion of the series, each child was tested individually for his role expectations of the status held by the central character in each drama, a taxi-driver. The test was a completion test, presented as an achievement test, in which the child was to give outcomes to various news

stories from current editions of the local newspaper. Subjects who did not understand the reality-reporting function of a newspaper were dropped from the analysis, for the children's responses to the newspaper test were taken as representing their reality expectations.

To permit two approaches to testing the hypothesis, two modes of scoring the completions were employed. Both were reliable: judges showed 92.5 per cent agreement on one scoring technique and 95 per cent agreement on the other.

The two sets of scores were used to test two predictions derived from the research hypothesis. The first prediction was that the E children would include more aggression, relative to the C children, in their completions to news stories featuring taxi-drivers than in their completions to others. This was confirmed. The second prediction was that in their completions to the news stories featuring taxi-drivers, the E children would attribute more aggression to the taxi-drivers than would the C children. This prediction was confirmed for one news story but not for the other. The news story for which it was confirmed was one which replicated the initial section of one of the radio dramas which the subjects had heard. Thus, in general, the data lend empirical support to the research hypothesis.

The discussion of the results includes an interpretation of the results summarized above and refers to other questions to which answers must be found in experimentation before an over-all assessment can be made of the influence of violence in the mass media on children's reality expectations.

References

1. Baldwin, A. L. *Behavior and development in childhood.* New York: Dryden, 1955.
2. Faust, Margaret S. Father-child relations in story completions. In Lois M. Stolz *et al., Father relations of war-born children.* Stanford, Univer. Press, 1954. Pp. 244-264.
3. Krech, D., & Crutchfield, R. S. *Theory and problems of social psychology.* New York: McGraw-Hill, 1948.
4. Linton, R. *The cultural background of personality.* New York: Appleton-Century, 1945.
5. Maccoby, Eleanor E., Levin, H., & Selya, B. M. The effects of emotional arousal on the retention of film content: A failure to replicate. *J. abnorm. soc. Psychol.,* 1956, 53, 373-374.
6. McClelland, D. C. *Personality.* New York: Dryden, 1951.
7. Sarbin, T. R. Role theory. In G. Lindzey (Ed.), *Handbook of social psychology.* Vol. I. *Theory and method.* Cambridge, Mass.: Addison-Wesley, 1954. Pp. 223-258.

8. SIEGEL, ALBERTA E. Film-mediated fantasy aggression and strength of aggressive drive. *Child Develpm.*, 1956, 27, 365-378.

9. THURSTONE, THELMA G., & THURSTONE, L. L. *Primary mental abilities test for ages 5 to 7.* Chicago: Science Research Associates, 1953.

10. *Television and juvenile delinquency: Report of the Committee on the Judiciary, United States Senate.* Washington: U. S. Government Printing Office, 1956.

GROUP CONFORMITY AND PERCEPTIONS OF PARENTS

PAUL HENRY MUSSEN
University of California, Berkeley

JEROME KAGAN
Fels Research Institute, Antioch College

A number of studies indicate that the tendency to yield to group opinion rather than to maintain independent judgment in experimental conformity situations is related to important personality variables (2, 3, 6). Thus, it has been found that college males who show a strong tendency to conform have a greater incidence of TAT dependency themes (*n* Succorance) than their less conforming peers, suggesting an essentially passive-dependent attitude toward adult sources of nurturance and support (6). Crutchfield's extensive studies of this problem reveal that, compared with those who are highly independent, extreme conformists are more authoritarian; more submissive, compliant, and accepting of authority; less tolerant, less socially active, and less able to accept responsibility; more anxious, over-controlled, and inhibited. Judging from their responses to items in personality inventories, conformists as a group are more intolerant of ambiguity and tend to emphasize conventional, socially approved values more than those who maintain independence in their judgments (2, 3). In interviews, they gave highly idealized descriptions of their parents, while independents offered a more balanced picture of praise and criticism. In general, extreme conformists have "rigid, externally sanctioned, and inconsistent moralistic attitudes" and "disturbed, dejected, and distrustful attitudes toward other people" (2, p. 195).

These data on the personality differences between conformists and independents suggest that early familial experiences may have an important influence on the adult's reaction to group pressure. More specifically, the conformist's submissiveness—reflected in his behavior in experimental conformity situations—may stem from his parents' early overly rigid and punitive treatment which prevented him from developing independence of thought and action. The independent's behavior, on the other hand, may be a consequent of less severe parental discipline. If this is true, and if the individual's perceptions of his parents, as revealed in fantasy productions, are correlated with early familial experiences, it may be hypothesized that extreme conformists would be more likely than independents to describe parents in TAT stories as harsh, restrictive, and punitive.

CHILD DEVELOPMENT, Vol. 29, No. 1 (March, 1958)

Method

The subjects, 27 male undergraduates, first wrote stories in response to eight TAT cards from the Murray series (1, 3BM, 6BM, 7BM, 12M, 13B, 14, and 18BM) and then participated in the conformity experiment designed by Asch (1). The procedure for this experiment followed Asch's instructions. The subject was led into a small room with four other men. Unknown to the experimental subject, these other four men were paid "stooges" cooperating with the experimenter. When the five men were seated, the experimenter gave the following instructions:

> This is a task which involves the discrimination of lengths of lines. You see a pair of white cards in front. The card on the left shows a single line; the card on the right has three lines differing in length. They are numbered 1, 2, and 3 in order. One of the three lines on the card on the right is equal in length to the standard line on the card on the left. You will decide in each case which is the equal line. You will state your judgment in terms of the corresponding number, either one, two or three. There will be 12 such comparisons. Since the number of lines is few and the group small, I shall call upon each of you in turn to announce your judgment, which I shall record here on a prepared form. Please be as accurate as possible. Suppose we start at the right and proceed to the left.

The naive subject was always seated fourth from the right and was always the fourth person to call out his judgment. When all five subjects had stated aloud their judgments on one pair of cards, these two cards were removed and replaced by a new pair with new standards and comparison lines. This was repeated 12 times with 12 different pairs of cards. The measurements of the standard and comparison lines were identical with those used by Asch (1).

The four men cooperating with the experimenter had been previously instructed to announce correct answers on five trials (1, 2, 5, 8, and 11) and to announce incorrect answers on the other seven trials (3, 4, 6, 7, 9, 10, and 12). On each of these seven crucial trials, all four announced the same incorrect answer which then constituted a majority opinion.

Nine of the subjects, the high conformers, yielded to the incorrect group judgments on two or more of the crucial trials in the Asch situation, while 11 of them, the independents, never conformed to the incorrect group opinion. Seven subjects conformed only once and are not included in this study.

A clinical psychologist[1] analyzed the TAT protocols of the nine high conformers and the 11 independents, scoring each story for the presence or absence of parental punishment themes. Such a theme was considered present if parents in the story, individually or together, deprived, rejected, and/or physically or verbally punished their child. (Examples: "The boy

[1] The authors wish to acknowledge the assistance of Mr. Curtis Hardyk in this aspect of the study.

can't go out to play baseball because his mother is making him practice his violin lessons." "The mother is telling her son never to return home.")

The scoring was done "blind," i.e., the judge was not informed of the purpose of the study; hence, his TAT analysis could not have been biased by knowledge of the hypothesis. Since punishment was very precisely defined, scoring reliability was high. For 48 stories, scored independently by the judge and one of the authors (PM), the percentage of agreement was 93.

A *parental punishment* score was derived for each subject by simply counting the number of such themes in his stories.

Results and Discussion

The two types of data obtained, conformity or nonconformity in the Asch situation and parental punishment scores, were used to test the hypothesis. For this group of subjects, parental punishment themes were relatively infrequent, the range being zero to three. Eight subjects had scores of zero, while six had scores of one. The distribution was therefore dichotomized into negative (no parental punishment themes) and positive (score of one or more) groups.

If the hypothesis is valid, a greater proportion of extreme conformists than of independents should have positive parental punishment scores. Table 1 shows the frequency of positive and negative scores among conformists and independents. Fisher's exact test for a 2×2 contingency table (4) was applied to these data and yielded the highly significant value of p between .025 and .01.

Table 1

DISTRIBUTION OF POSITIVE AND NEGATIVE PARENTAL PUNISHMENT SCORES AMONG CONFORMISTS AND INDEPENDENTS

Group	Parental Punishment Scores	
	Positive	*Negative*
Conformists	8	1
Independents	4	7

Thus, the data support the hypothesis that conformity with the group is related to perceptions of parents, as reflected in fantasy productions, as harsh and punitive individuals. Assuming that these perceptions are the consequents of past experiences with parents rather than current views of authority in general, these findings suggest that the quality of early parent-child relationships may exert an important influence on the adult's conformity tendencies.

The data from Crutchfield's study, as well as from the present one, are consistent with the interpretation that conformity with the group may be a response which has been generalized from early-learned reactions toward feared parents. As Anna Freud has pointed out (5), a child may be able to reduce the anxieties stemming from parent-child relationships by *identification with the aggressor,* i.e., by imitating, and conforming with, the behavior of the parent he fears. Those who feel harshly treated by their parents probably have more conflicts with them and hence probably have more experiences in which conformity is rewarded, i.e., anxiety-reducing. Consequently, their conformity responses are likely to acquire relatively great habit strength and to be generalized to relations with other authorities, including the authority of group opinion. In this connection it is interesting to note that in Crutchfield's study conformists revealed marked attitudes of distrust toward others. Such attitudes may also be generalized responses stemming from original fear and distrust of the parents.

Summary

The purpose of the study was to relate conformity tendencies to perceptions of the parents as shown in fantasy. Twenty-seven male college students wrote stories in response to eight TAT cards and then were individually observed in the Asch conformity situation.

A significantly greater proportion of extreme conformists than of independents perceived their parents as harsh, punitive, restrictive, and rejecting. These data, like those from previous studies, suggest that conformity tendencies are manifestations of basic personality structure and are influenced by early parent-child relations. The data may also be consistent with the interpretation that extreme conformity in the experimental situation represents a generalization from early-learned conformity behavior.

References

1. Asch, S. E. *Social psychology.* New York: Prentice-Hall, 1952.
2. Crutchfield, R. S. Conformity and character. *Amer. Psychol.,* 1955, 10, 191-198.
3. Crutchfield, R. S. Recent studies in conformity behavior. Paper given at Twelfth Regional Meeting of Veterans Administration Psychologists, California Academy of Sciences, San Francisco, November 17, 1955.
4. Fisher, R. A. *Statistical methods for research workers.* (7th Ed.) Edinburgh: Oliver & Boyd, 1938.
5. Freud, Anna. *The ego and the mechanisms of defense.* New York: International Universities Press, 1946.
6. Kagan, J., & Mussen, P. Dependency themes on the TAT and group conformity. *J. consult. Psychol.,* 1956, 20, 29-32.

THE BEHAVIOR-INFERRED MOTIVATIONS OF LATE- AND EARLY-MATURING BOYS

PAUL HENRY MUSSEN *and* MARY COVER JONES
University of California, Berkeley

A number of studies indicate that the adolescent's physical status, mediated by the sociopsychological environment, may exert profound and lasting influences on his personality (4, 5, 6). In our culture, physically retarded adolescent boys are in a disadvantageous competitive position in athletic activities and are likely to be regarded and treated as immature by both adults and peers (5). Hence, they are more likely than their physically accelerated peers to be personally and socially maladjusted during adolescence. Analysis of the TAT responses of early- and late-maturing boys about 17 years of age, subjects in the Adolescent Growth Study (2, 3), revealed that significantly more of the latter group had negative self-concepts, profound feelings of rejection by others, strong affiliative needs (especially for heterosexual affiliation), prolonged dependency needs, and rebellious attitudes toward parents. The psychological picture of the physically accelerated boys, on the other hand, seemed to be much more favorable. As a group, they appeared to have acquired more self-confidence, were less dependent and in need of help, and were more capable of playing an adult male role in interpersonal relations. Contrary to what had been predicted, the two groups did not differ in needs for achievement and recognition, and there was some evidence that early-maturers had stronger aggressive needs than late-maturers (6).

The motives reflected in fantasy productions are presumably covert and not readily apparent in ordinary manifest behavior. It has been demonstrated that, in many cases, manifest behavior and fantasy may be independent or even negatively related (7, 8). Hence, inferences about motivation based on the individual's overt behavior may be quite different from those made on the basis of his fantasy productions.

In the present study, we shall examine the *behavior*-inferred ratings of motivation assigned to early- and late-maturing boys. Under Else Frenkel-Brunswik's direction, almost all subjects in the Adolescent Growth Study were rated on nine drives about a year after they had graduated from high school (1). The strength of each drive was inferred from behavior

observed in a wide variety of situations. " 'Intuitive' judgments about motivational tendencies were obtained in the form of standardized ratings from a number of judges well acquainted with the adolescents" (1, p. 134).

The nine drives, selected from Murray's list of needs, were briefly described as follows:

1. *Drive for Autonomy. Striving for independence and freedom;* desire to be free from social ties, to shake off influence, coercion, and restraint; relatively little care for conventions and group ideology; tendency to act as one pleases.

2. *Drive for Social Ties, Social Acceptance. Desire to be generally well-liked;* to conform to custom, to join groups, to live sociably, to be accepted by a group in any form, to make contacts.

3. *Drive for Achievement. Desire to attain a high standard of objective accomplishments;* to increase self-regard by successful exercise of talent, to select hard tasks; high aspiration level. (Rating scale starts from "no desire to accomplish something outstanding" and ends with "excessive demands on himself.")

4. *Drive for Recognition. Desire to excite praise and commendation, to demand respect, social approval and prestige, honors and fame.*

5. *Drive for Abasement. Tendency to self-depreciation, self-blame or belittlement;* to submit passively to external forces, to accept injury, blame, criticism, punishment; tendency to become resigned to fate, to admit inferiority and defeat, to confess, to seek punishment and misfortune; masochistic tendency.

6. *Drive for Aggression. Desire to deprive others* by belittling, attacking, ridiculing, depreciating.

7. *Drive for Succorance. Desire for support from outside;* from people, institutions, or supernatural agencies.

8. *Drive for Control (Dominance). Desire to control one's human environment,* by suggestion, by persuasion or command.

9. *Drive for Escape. Tendency to escape all unpleasant situations;* to avoid blame, hardship, etc., to project own failures on others or on circumstances; to gain immediate pleasure with inability to postpone pleasure; use of fantasy, etc. (1, pp. 141-142).

"The rating scale was used as a means of summarizing what the judges 'privately' think, or what their hypothesis was, about each subject, or what they otherwise would have used only in writing up an interpretative case study of individuals whom they observed over a long period of time" (1, p. 144). "Though ultimately referred to observed behavior, reference to the underlying motivations was established as the result of a complex process of inference utilizing more subtle indirect cues together with gross features of behavior" (1, p. 261).

Since the drive ratings were found to be significantly correlated with many ratings of social responses and personality traits, Frenkel-Brunswik concluded that they "helped to organize the . . . data on overt behavior observed in social situations" (1, p. 261). "One drive variable [may circumscribe] a family of alternative [behavioral] manifestations unrelated to each other. The meaning of the drive concept emerges in terms of fam-

ilies of divergent manifestations held together dynamically or genotypically, though often not phenotypically" (1, p. 262).

In view of the interrelationships discovered, it may be assumed that these drive ratings refer to "a level of personality which stands behind the surface of overtly displayed social techniques" (1, p. 144), but which is probably not generally as remote from behavior as drives reflected in responses to projective tests. Comparison of late- and early-maturers' motivational ratings may, therefore, provide additional information about the factors underlying behavioral differences between the two groups. The objective meanings of these ratings may be derived from examination of the variables significantly correlated with them.

Subjects and Procedure

The 34 subjects of this study, part of a normal sample of 90 boys who were members of the Adolescent Growth Study (2, 3), were drawn from the two extremes of a distribution of rate of physical maturing. Sixteen of them had been among the most consistently accelerated throughout the adolescent period; the other 18 had been among the most consistently retarded.[1]

Three members of the Adolescent Growth Study staff who were well acquainted with the subjects over a period of years served as judges. They independently rated each subject on each of the nine drives listed above. In making the ratings, the judges were asked "to forget about manifest behavior and to group the children according to assumed motivation rather than according to similarities of displayed techniques" (1, p. 139).

Each drive was rated on a five-point scale, a rating of one representing the highest degree, a rating of five representing almost complete absence of the drive in question. In the present study, the three judges' ratings on each drive were averaged to obtain the individual's motivational rating.[2]

Results

Frequency distributions of the ratings of all subjects were made for all the drives. Each distribution was then dichotomized at the point which most nearly enabled the placing of half of the 34 subjects above, and half of them below, the dividing point. Subjects having scores above this point

[1] The present sample includes all those subjects among the 19 most consistently accelerated and 19 most consistently retarded boys in the Adolescent Growth Study who were assigned motivational ratings. Twenty-nine of them were also in Jones and Bayley's (5) group and 31 in Mussen and Jones' (6).

[2] The interrater agreements on the nine drive ratings assigned to all the subjects of the Adolescent Growth Study are fully reported by Frenkel-Brunswik. The average reliabilities, corrected by means of the Spearman-Brown formula, were between .70 and .80, with the exceptions of those for Autonomy and Succorance which were .55 and .63 respectively (1, p. 100).

were considered *high* in this particular drive; those with scores below this point were considered *low* in it. Chi square tests were used to ascertain whether or not high ratings in certain drives were more characteristic of one group (late- or early-maturers) than of the other.

Table 1 shows the number of late- and early-maturers with high ratings in each of the drives, the chi square values obtained, and the levels of significance.

TABLE 1

NUMBER OF EARLY- AND LATE-MATURERS ASSIGNED HIGH DRIVE RATINGS

	Number Rated High			
Drive	*Early-Maturers* ($N = 16$)	*Late-Maturers* ($N = 18$)	*Chi Square*	*p*
1. Autonomy	7	10	.47	ns
2. Social Ties	6	13	4.37	.02–.05
3. Achievement	10	9	.53	ns
4. Recognition	7	9	.15	ns
5. Abasement	9	8	.47	ns
6. Aggression	5	12	4.27	.02–.05
7. Succorance	5	9	1.30	ns
8. Control	9	9	.13	ns
9. Escape	9	11	.09	ns

As the table shows, the two groups were significantly different in two of the nine drives. A greater proportion of late- than of early-maturing boys were rated high in *Drive for Social Ties; Social Acceptance* and *Drive for Aggression.*

While these drives overlap to some extent, they are by no means identical. The correlation between them, based on the entire Adolescent Growth Study male population, was .48 (1). It should be noted, however, that the early- and late-maturing groups did *not* differ significantly in other drives (*Drives for Recognition, Control, and Escape*) that were even more highly correlated with *Drive for Aggression* and/or *Drive for Social Ties* in the entire sample of adolescent boys (1).

It is of course possible that the independent (physical status) and dependent (motivational ratings) variables were not completely independent, i.e., that in evaluating underlying drives the judges were influenced by their own impressions of the subject's rate of physical maturation. However, it seems unlikely that this factor had any significant effects, since the judges were specifically instructed to rate "assumed motivations" rather than more superficial characteristics. Moreover, if the judges had

been considerably influenced by the subjects' rates of maturation, the two groups might be expected to differ significantly on more than two of the nine drive ratings.

Discussion

The finding that late-maturers appear to have relatively high motivation for social affiliation and social acceptance is entirely consistent with Jones and Bayley's data which showed that observers rated these boys as higher in sociability, social initiative and eagerness (5). Moreover, the TAT protocols of the late-maturers revealed that these boys have strong needs for affiliation, particularly with the opposite sex (6).

On the basis of evidence reported in earlier studies, it has been hypothesized that the late-maturers' emphasis on social activity and their social initiative and participation are largely of an attention-getting, compensatory nature (5, 6). Examination of the overt social behavior and personality trait variables correlated with *Drive for Social Acceptance* lends support to this hypothesis. For example, high ratings in this drive tended to be associated with high ratings in *energy output, interest in opposite sex, social participation, social self-confidence, social stimulus value, popularity, self-assertiveness* (bossiness), *exuberance, orientation toward the opposite sex*[3] (1). These are characteristics which indicate expressiveness, energy, eagerness and sociability, a constellation labelled "overt social activity" by Frenkel-Brunswik. At the same time, however, high *Drive for Social Ties* was associated with many traits considered to be indicative of emotional maladjustment: *attention-seeking, affectation, dependence on approval, excitedness, tenseness, impulsiveness,* and *exploitiveness* (1). Hence, it may be inferred that, while late-maturers often reveal high motivation for social affiliation and a great deal of social activity, their social techniques are often childish and affected. Moreover, their high social drives may be based on general insecurity (reflected in *tenseness* and *impulsiveness*) and basic feelings of dependency (indicated by *dependence on others* and *exploitiveness*). The last conclusion is consistent with the finding that the TAT protocols of late-maturers revealed strong underlying dependency needs (6).

The relatively high aggressive motivation ratings assigned to the late-maturers may also be attributed, at least in part, to their basic insecurities, feelings of rejection, and intense dependency needs. Frenkel-Brunswik's data showed that *Drive for Aggression* had considerably higher correlations with emotionally maladjusted behavior than with "overt social activity." Among the manifestations of the latter, only four (*energy output, interest in opposite sex, social participation,* and *social self-confidence*) were significantly correlated with *Drive for Aggression*. On the other hand, 10 indi-

[3] The correlations between the drive ratings and social behavior are fully reported in Frenkel-Brunswik's monograph (1, p. 186-7).

cators of emotional maladjustment (*attention-seeking, affectation, dependence on approval, excitedness, tenseness, impulsiveness, frequent mood swings, selfishness, irresponsibility,* and *exploitiveness*) were associated with high ratings in aggressive motivation. Apparently, the personal and social maladjustments of the late-maturers are often reflected in overt behavior which trained observers are likely to interpret as stemming from strong underlying aggressive motivations.

This finding is particularly interesting in view of the fact that late-maturers had relatively *low* fantasy (TAT) aggression scores, while the early-maturers had *high* scores in this variable (6). The present data do not enable us to specify the determinants of this discrepancy between behavior-inferred aggressive ratings and TAT aggression scores. Some of it may be due to the middle-class status of most of our subjects. It has been suggested that in this social class cultural prohibitions prevent the overt gratification of aggressive needs and thereby increase their intensity in the individual's fantasies (7, 8). If this is true, it may be hypothesized that the early-maturers, having identified with a mature male role, have incorporated these prohibitions and are less likely to behave aggressively. As a result, they are more likely to express their aggressive motivations in fantasy. On the other hand, the late-maturers, being defiant of authority (6), may not accept the middle-class standards of behavior so readily. Consequently, they may express aggression more freely and overtly and, therefore, may reveal less aggressive motivation in their fantasy productions.

It may be concluded that these group differences in motivational ratings add further support to the general hypothesis that the late-maturing adolescent is more likely to be personally and socially maladjusted than his early-maturing peer. Apparently in our culture, the physically retarded boy is more likely to encounter a sociopsychological environment which may have adverse effects on his personality development. The early-maturer's experiences, on the other hand, seem to be more conducive to good psychological adjustment.

Summary

In this investigation, observers' ratings of 18 late-maturing and 16 early-maturing adolescent boys on nine drives were compared. The results showed that high drives for social acceptance and for aggression are more chararacteristic of the physically retarded than of the physically accelerated.

The meaning of group differences in these ratings may be clarified by examination of the social behavior and personality correlates of the two drives. These suggest that the late-maturer's high social drives may stem from feelings of insecurity and dependence and are often manifested in childish, affected, attention-getting social techniques. Moreover, high aggressive drives tended to be associated with social behavior and personality characteristics indicative of social and emotional maladjustment. Thus,

among the physically retarded, the strength of these drives may be partially attributable to underlying feelings of inadequacy, rejection, and dependence.

In general, the data of this investigation support the findings of earlier studies which showed that, among boys, physical retardation may have adverse effects on personality. Physical acceleration, on the other hand, may be conducive to better social and psychological adjustment.

References

1. Frenkel-Brunswik, Else. Motivation and behavior. *Genet. Psychol. Monogr.*, 1942, 26, 121-265.
2. Jones, H. E. Observational methods in the study of individual development. *J. consult. Psychol.*, 1940, 4, 234-238.
3. Jones, H. E. *Development in adolescence.* New York: Appleton-Century, 1943.
4. Jones, M. C. The later careers of boys who were early- or late-maturing. *Child Develpm.*, 1957, 28, 113-128.
5. Jones, M. C., & Bayley, N. Physical maturing among boys as related to behavior. *J. educ. Psychol.*, 1950, 41, 129-148.
6. Mussen, P. H., & Jones, M. C. Self-conceptions, motivations, and interpersonal attitudes of late- and early-maturing boys. *Child Develpm.*, 1957, 28, 243-256.
7. Mussen, P. H., & Naylor, H. K. The relationship between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, 49, 235-240.
8. Sanford, R. N., Adkins, N. M., Miller, R. B., Cobb, E. A., *et al.* Physique, personality, and scholarship: a cooperative study of school children. *Monogr. Soc. Res. Child Develpm.*, 1943, 8, No. 1.

FACTORS INFLUENCING INFORMAL GROUPS OF ADOLESCENTS

HAROLD R. PHELPS *and* JOHN E. HORROCKS
Institute for Child Development and Family Life
The Ohio State University

It has been the purpose of this study to investigate attitudes and activities that lead to the formation of informal adolescent groups. A large portion of adolescent social life is organized around small informal groups. Such groups are the proving ground for the teen-ager's widening perception of his social role and the exploration of techniques to implement that role. The informal group plays a much more important part in the life of the adolescent than do formally organized groups and may be thought of as a focal area of experience in the process of coming of age. Observation of informal adolescent groups leads to the conclusion that while the group is more or less ephemeral in its membership and continuity it does center around certain commonalities possessed by its members. Membership is usually exclusive in that some youth are excluded or are made to feel uncomfortable when participating in the group's activities. An understanding of adolescent behavior would be advanced by an analysis of the kinds of activities and attitudes important in the formation and perpetuation of such groups with particular reference to those things which contrast one group to another as well as to the commonalities which bind them all together as part of the larger adolescent age peer society.

Previous studies of group activities of youth have characteristically centered around a normative survey of activities participated in on the basis of such categories as age, grade, sex, intelligence, and socioeconomic status. Important as these categories have proved to be it would seem that more meaningful classifications could be developed from the interrelationships existing among data concerning the behavior of groups. Age, sex, and place of residence together with similar categories should be considered simply as other variables along with such facts as whether a subject's group plays ball, goes fishing together, rejects people who date, or makes use of a certain playground in their community. Analysis of activity and

CHILD DEVELOPMENT, Vol. 29, No. 1 (March, 1958)

attitude data against such major categories as sex rests on the assumption that a casual relationship exists between the categories and the data analyzed, with the major categories playing the role of independent variables. With the exception of age and sex there is a growing body of data which would indicate that the assumed causal relations between the categories and various activities and attitudes may be reciprocal (1). As an example, some persons may remain in a lower socioeconomic classification merely because they retain specific patterns of behavior and attitudes. There is also ample evidence to indicate that measured intelligence may be influenced by the opportunity for certain social experiences (2, 3). Thus, it would seem a more fruitful procedure to place age, grade, sex, socioeconomic status, etc., within the context of activities and attitudes involved in belonging to a group and allow them to find their places within categories on the basis of their relationship with all the data involved. If any variable emerges as a category or dominates a category, it has been validated. The present study has adapted such a procedure.

Subjects

The 200 subjects for the present study were obtained from the 800 children enrolled in grades 7 through 12 in the public and parochial schools of a central Ohio community of approximately 9,500 persons. The 200 comprised a sample drawn randomly from among 726 children, stratified according to grade, who filled out the questionnaire on which this study was partly based and for whom comprehensive personal information was available. Most of the youth in the study lived in the city although a few commuted to school from the surrounding rural area.

Procedure

All 800 children in grades 7 through 12 were administered a questionnaire entitled *Recreation in Your Community*. Of the 800 questionnaires given to students, 726 were returned. The questionnaire was omnibus in type. Six major areas of information were investigated as follows: (a) identification of the children by name, age, sex, grade in school, address, and other data of similar nature; (b) location and identification of informal adolescent groups by asking each subject to give names and other identifying characteristics of his favorite group or groups; (c) information concerning group attitudes, such as inclusiveness and exclusiveness, group ideals, and estimated group reactions toward samples of typical adolescent behavior; (d) measurement of the subject's feelings toward his own group and toward other groups with which he was acquainted; (e) naming of the activities and meeting places of the groups identified; (f) measurement of each subject's satisfaction with his groups and their activities.

In order to avoid the subjectivity of compiling items for a questionnaire by trying to anticipate possible responses that subjects might make, a series of both structured and open-ended interviews were conducted. Interviews about group membership and related matters were held with 100 subjects selected at random from the children in grades 7 through 12. An analysis of the interviews together with an extensive survey and field observation of the community in which the study was conducted made it possible to construct appropriate and information-producing items about informal adolescent groups and their activities in that community. A preliminary form of the questionnaire was designed and administered to a small group of subjects as a means of checking for ambiguity, interest appeal, and answerability of the questions included. It was felt that with subjects of such a wide range of ages and backgrounds it was particularly important to employ language capable of being understood by the youngest class of respondents.

In its final form the questionnaire[1] consisted of six sections named as follows: (a) What you would like to do for recreation, (b) Your group, (c) How do you feel about your group? (d) Whom you like in your group, (e) What your group does, and (f) Where your group goes.

Wherever possible check type questions were used. Considerable effort was made to make the questionnaire interesting to those who were asked to fill it out. Among other devices intended to achieve interest, the questionnaire was illustrated with line drawings. Every effort was made to refer to the respondent as an individual and to make him feel that the study was really interested in, and highly valued, his opinions. This was done, in part, by directing the questions to the individual, by using personal pronouns, and by heading each group of questions by personal reference statements such as "Your Group." With the cooperation of community authorities it was also possible to promise children who answered the questionnaire that recreational needs revealed by their answers would, whenever possible, be incorporated in the recreational programs of the schools and other community agencies.[2] Questionnaires were filled in and left in the local school offices in envelopes marked "confidential." Subjects were assured that their questionnaires would be available to no one in their community.

In addition to the information obtained by the questionnaire the following data on each subject were available: (a) socioeconomic status based on Warner's Index of Status Characteristics (4), and (b) intelligence as recorded in the local school records.

---

[1] A detailed discussion of the questionnaire and its construction will be presented in a forthcoming article by Horrocks and Bowlus.

[2] At the end of the study, through cooperation with the City Recreation Commission and the public schools, a number of the children's desires as revealed by the questionnaire were taken into consideration in planning the community's recreational program.

Analysis of the Data

Of major concern in designing this study was the selection of a technique of statistical analysis which would allow all data equal opportunity of influencing the final outcome of the study. Multiple factor analysis appeared the best technique to meet this criterion. However, from the questionnaire and socioeconomic data of the present study there were 168 items to be treated as separate variables. This would have made it necessary to start the analysis with 14,028 correlation coefficients in the factor matrix. Under the circumstances it was decided to use the Wherry-Gaylord (5) method of obtaining oblique factors by iteration utilizing easily computed tetrachoric correlation coefficients.[3]

One hundred and fourteen variables concerning attitudes and activities were subjected to the Wherry-Gaylord method of analysis. To make computation of the tetrachoric correlations simpler, 200 cases of the original 726 were selected randomly, stratified according to grade in school. In this way simple frequencies could be read as percentages in the calculation of the tetrachoric coefficients of correlation. From this analysis emerged 10 oblique factors. Loadings on each of the 10 oblique factors were computed for the variables which had not been included in the first analysis.

By means of a modification[4] of the Doolittle solution for solving simultaneous equations, the oblique factors were rotated to orthogonality. Loadings for all variables on the newly obtained orthogonal factors were computed. These orthogonal factors were rotated visually, attempting (a) to maximize the loadings of variables on the factor to which they appeared to belong, and (b) to minimize their loadings on all other factors. Ultimately, however, the criterion for rotation was meaningfulness. It was believed that mathematically neater categories are of little value if they do not make sense psychologically.

Visual rotation of 168 variables would have been next to impossible. As an alternative, a sample of 47 variables was selected for rotation. An attempt was made to represent evenly all the factors (a) with variables which had a high loading on a particular factor and few loadings on other factors, and (b) with variables which had relatively heavy loadings on several factors. The latter representation was decided upon to make certain that overlap between the factors would be taken adequately into account. After the factors were rotated utilizing the sample variables, new factor loadings resulting from the rotations were computed for all variables. It was then found that the solution could be improved with three additional rotations. In all, 45 rotations were made. An interpretation of the factors rotated to psychological meaningfulness follows in the next section.

---

[3] The tables for this study were so extensive that it was impractical to include them in the present publication. They have been mimeographed and may be obtained by writing either of the authors.

[4] Obtained from Dr. Robert J. Wherry, The Ohio State University.

Interpretation of Factors

Ten factors emerged from the factor analysis and were labeled I through X. Although loadings of .20 or more were significant at the 1 per cent level of confidence, none below .25 was included in the interpretation. Loadings lower than .25 were omitted because it was believed that their inclusion would interfere with interpretation to a greater degree than the amount they might add to the meaning of a factor. Since all variables were treated as dichotomies their names were changed in the interpretation so that most loadings on the factor being considered were positive. Such positive loadings made interpretation of the factors less complicated and more meaningful.

Factor I

Factor I depended to a very considerable extent on having automobile transportation available, being able to get around freely in the community, and going a considerable distance outside the small city in which the present study was carried out. Adolescents scoring high on Factor I liked riding around in cars (.75),[5] preferred people who had cars (.56), frequently went to a large city (.62) about 25 miles distant, and expressed a desire to leave town to have a good time (.52, .54).[6] A water gym (.36) with picnic grounds attached where there was considerably less supervision than at the city swimming pool was patronized. This recreational facility was located about three miles from the small city. Walking around town (.48) and window shopping (.62) were popular activities. The bowling alley (.69) was a favorite meeting place.

Much activity involved going to various establishments serving food. Enjoyment of eating at restaurants (.65) was expressed, but this was by no means the only reason for patronizing them since sitting or talking in restaurants (.49) was frequently listed as an activity. Drive-in type restaurants were preferred (.79, .70), and one about three quarters of a mile outside of town (.53) was frequently mentioned. Four ordinary eating places (.57, 51, .51, .42) and several dairy bars (.53, .43, .42, .37) were also named.

Having dates (.70) was important to young people scoring high on Factor I. They preferred people in their group who liked to date (.57) and were not likely to disapprove of people in the group who "neck" (.29) and "park" (.29). They liked to go to a drive-in theatre (.58) about one and a half miles outside the city. They also liked to go to regular movie theatres (.43). Most often the favorite group named was heterosexual in its make-up (.64). The groups sometimes went to an old, abandoned slate

[5] Numbers with decimal points in front of them enclosed in parenthesis are factor loadings of the variable immediately preceding them on the factor being discussed.

[6] Two items expressed essentially the same idea.

pit (.34). Most frequent time of meeting was on Sundays (.38) and at night (.28).

Activities appealing to adolescents scoring high on Factor I were playing cards (.26), listening to the radio (.42), watching television (.29), and having parties (.26). They liked to play baseball (.32). It did not seem to matter much to them if members of the group smoked (.40) or were not of the same age (.28), but they tended to prefer that members of the group not be colored (.28).

Young people influenced by Factor I were likely to be 17 years of age or older (.57) and to be no younger than 14 (.56). They were for the most part members of grades 11 and 12 (.53), and there was some tendency for seventh graders (.38) *not* to show the influence of the factor. There was some evidence that adolescents of IQ 111 or better (.28) whose fathers were in the upper occupations (.28) and whose families were of upper socioeconomic status (.26) were *not* likely to name a group influenced by the factor.

The satisfaction of heterosexual interests and a desire for peer-group activities not associated with the home were important pressures in this factor. None of the activities named was home-centered or home-influenced to any great degree. Associated with extra-home interests appeared a need for ranging about and going some distance from the community for social participation. It seems reasonable to assume that the forces underlying Factor I are a combination of pressures leading to the assumption of an adult role. Certainly, getting away from the influence of the home and ranging over territory farther and farther away from the home and community are important aspects of becoming an adult in the American culture. Of equal importance in becoming an adult are the physiological and social pressures which culminate in sexual mating.

It is of interest that lower socioeconomic status and lower intelligence were somewhat associated with Factor I. One expects adolescents whose parents have limited financial resources to assume the adult role sooner than those whose parents are more comfortably fixed. Previous studies have established that earlier emancipation from home, earlier assumption of the adult role, and greater sex activity are associated to some degree with lower socioeconomic status.

It is proposed that the present factor be described as a: *Pattern of pressures leading to assumption of the adult role, emancipation from the home, and satisfaction of heterosexual interests.*

Factor II

One variable stood out markedly in this factor. People who sometimes cheat to get by (.86) were not acceptable to groups influenced by Factor II. There were many behaviors which were thought to be inappropriate for group members. People who gamble (.57), smoke (.57), tell off-color jokes (.44), "park" (.44), "neck" (.43), ride motorcycles (.40), are noisy (.38),

swear (.32), and race cars (.31) were not acceptable. It was preferred that members of the group be persons of about the same age (.30), who have a good reputation (.36), nearly always agree (.34), and think good manners are important (.32). They thought that it was important for members to go to church (.30) but it was *not* important whether they went to churches other than their own (.30). Playing cards (.28) tended to be a rejected activity. There was a tendency to reject those who lived in the best part of town (.28).

There was much evidence that both home and school were forces in producing Factor II. Adolescents were preferred in the group who did not get in trouble with teachers (.57), did not quit school (.42), and were good students (.33). In relation to family matters, they expressed a preference for people who "stick-up" for their families (.48) and usually mind their parents (.28). The groups frequently met at places which were either part of the school or sponsored by the school. The assembly at high school (.54), noon movies at school (.52), and the high school stadium (.27) were favored places to go. They went to school athletic events (.34) together. One of the elementary school grounds (.42) located in a better part of town was a meeting place, and they patronized a recreation center (.31) sponsored by the school and located in one of the elementary school buildings.

Groups influenced by Factor II frequently went to the city swimming pool (.53) and liked to go swimming together (.28). They went together to movie theatres (.51), but tended to avoid the drive-in theatre (.31) so popular with adolescent groups identified with Factor I. Bicycle riding (.45) and hiking (.27) were favored physical activities, but it was likely that groups influenced by the factor would not play baseball (.25). They sometimes watched television (.26).

They patronized a hamburger "joint" (.40) in the business district, but tended to avoid a dairy bar (.39) also avoided by adolescent groups influenced by Factors III and V. One of the best restaurants in town (.26) most popular with adolescents of high socioeconomic status influenced by Factor V and avoided by those of low socioeconomic status reacting to the pressures of Factor IV was a meeting place for the Factor II group. They expressed a mild interest in sitting in restaurants and talking (.26).

Adolescents naming groups scoring high on Factor II were likely to to be in the eighth grade (.38) and to come from families who obtained high, composite, socioeconomic ratings (.31). Their fathers tended to be employed in an upper socioeconomic occupation (.26).

With the high concentration of variables indicating disapproval for behavior and attitudes which are generally considered to be socially unacceptable, it may be hypothesized that Factor II has to do with a moralistic point of view or a need to conform to certain social standards. The fact that many of the activities were connected with school, that there was a moderately high loading on identification with persons and activities having upper socioeconomic status, and an emphasis upon family loyalty

suggests that the pressures behind this factor may be largely those induced by an upper socioeconomic home and school companions from homes of such status. One inconsistency was the appearance of a loading, however small, on the rejection variable for those who live in the best part of town. Such inconsistency might possibly be explained as a simple stereotype and expression of the belief that one should not overtly choose one's friends on the basis of their living in the best part of town. Although loadings on this factor were relatively low on variables having to do with church, there is some reason to believe that Factor II may be fostered by religious background since many of the values held by young people scoring high in this factor are characteristic of numerous religious groups.

Factor II is designated as follows: *Pattern of pressures assuming the form of a moral code approved by the school and upper socioeconomic home.*

Factor III

Very prominent in Factor III were variables having to do with going to business establishments where ice cream and other dairy bar products might be purchased. Going to three dairy parlors had loadings in the .60's. Two of the dairy bars had no provision for remaining after the purchase had been made. For some reason which the writers were unable to determine, one dairy bar was rejected with a negative loading of .33.

Noticeable were variables related to school life such as going to school affairs (.46), studying together (.52), going to the public library (.48), to school lunch room (.39), and assembly at the high school (.39).

There were many activities included in the factor. Those with highest loadings were having picnics (.61), listening to the radio (.59) and records (.57), bicycle riding (.57), making candy (.56), playing cards (.59), roller skating (.55), and playing volley ball (.52). Having parties, watching television, walking around town, swimming and tennis had loadings in the .40's. Only one activity was rejected and that was playing baseball with a negative loading of .32.

Groups scoring high on Factor III were likely to have seven or more adolescents in them (.35) and tended to be named more often by girls (.37). Young people naming groups high in this factor were likely to be in the eighth grade (.27), preferred the group they named (.28), and have fathers in the middle occupational classification (.26). Whether or not the person lived in the best part of town (.29) or wanted to quit school (.36) did not seem to matter.

It is very likely that this factor represents a cluster of activities around which some informal groups are organized. Most of the activities are ones which would generally be approved by teachers and parents. In contrast with Factor I all of the activities are ones which can be carried on within a relatively short radius of the home. In fact, all could be done within the small community in which the study was carried out.

Perhaps Factor III is the result of supervision exercised by the family forcing the adolescent to engage in activities which may be carried out not too far distant from home or under the surveillance of responsible adults. This would appear to be a logical conclusion particularly since girls who are supervised much more closely in our culture had a relatively large loading (.37) on this factor. It is therefore proposed that Factor III be named: *Non-emancipated, home, school, and community centered activity pattern.*

Factor IV

Adolescents preferring a group characterized by behaviors and characteristics associated with Factor IV were most likely to be of low socioeconomic status. Their families very often received a composite, socioeconomic rating (.38) placing them in the lowest category. They lived in the poorest housing (.40) located in the poorest sections of the city (.31). It was quite unlikely that any of their fathers were to be found in occupations rated among those in the highest category (.41) utilized in the present study. There was a trend for them to have an IQ on group tests of intelligence less than 89 (.25).

They appeared to make discriminations concerning group membership along socioeconomic lines. It was preferred that group members' fathers be laborers (.89) and they were likely to reject people who lived in the best part of town (.69) and had more money than most (.50). Persons who lived in the same neighborhood (.75) and were about the same age (.34) were preferred. They liked people who had high aspirations expressing a preference for those who wanted to go to college (.74) and worked after school (.60). Strength was important to them (.84).

It was important that group members go to church a lot (.50) and go to other churches (.85). What was meant by going to other churches is difficult to understand. Perhaps it was not particularly important what churches their group members attended, but that they go a lot to whatever church they preferred.

They thought that it was important that members of their group be in early at night (.56). Some value was placed on good looks (.29) and having a car (.25). People who were always kidding (.55) and nearly always agreed (.32) were liked. Colored young people were sometimes acceptable in their groups (.29). Frequently they made candy (.50) when they were together and went to an abandoned slate quarry (.43) located in a poorer section of the city. Roller skating (.41) was a popular activity, but there was a tendency for them not to go to movie theatres (.30) as a group.

One restaurant (.38) of very poor appearance, located in the business district, was popular. It was also popular with the adolescents of Factor I. Three dairy bars (.33, .30, .28) were sometimes patronized, one of which was popular with adolescents of Factors I and III, but avoided by those influenced by Factor V. One of the drive-in type was also popular with

people influenced by Factor I. The remaining dairy-bar was popular in both Factors I and III.

Some restaurants were avoided. One (.40) of very modern appearance located in the business district and specializing in fancy candies was not patronized by young people influenced by this factor but was found desirable by the young people of Factors I, III, and V. The restaurant (.34) which was possibly the finest in the city and very popular with the upper socioeconomic adolescents of Factor V was also avoided.

Without question Factor IV is the result of socioeconomic differences which are reflected in the nature of the activities and attitudes finding their places in this category. There appeared to be identification with their socioeconomic peers and a rejection of others in more fortunate circumstances. Perhaps they sensed the social and economic differences between themselves and other adolescents and were trying to make up for them by having higher ideals as manifested by their attitude toward church attendance and their somewhat unrealistic aspiration of going to college. Lack of financial means must not have been the sole determiner of the kinds of activities they participated in for they were able to patronize several restaurants and dairy bars. It was significant, however, that they avoided to some extent places popular with young people of upper socioeconomic status.

It is proposed that Factor IV be considered a: *Pattern of activities and social values deriving from a very low socioeconomic status.*

Factor V

Groups influenced by Factor V were more often named by girls (.41) who did *not* live in the poorer sections of the community (.48) or whose fathers were *not* engaged in low status occupations (.43). They were more likely to have fathers in the highest status occupations (.34) and to live in the best part of town (.32). As a matter of fact, several of the socioeconomic variables had loadings ranging from .26 to .48, rather clearly indicating that adolescents scoring high on Factor V tended to come from upper socioeconomic homes and *not* from the lowest group in the community. There were no significant loadings on the middle group.

Adolescents scoring high on Factor V named groups with seven or more in them (.56). The activities most favored were having slumber parties (.78), going to two higher quality restaurants downtown (.75, .73), and one about three-fourths mile outside the community (.58) which required some kind of transportation to reach it. In general, they liked to eat (.39) and talk (.34) in restaurants but were likely to shun certain establishments serving dairy products (.40, .30, .25) which were popular with groups scoring high in Factors I and III.

They liked to dance (.53), make candy (.41), listen to records (.40), have parties (.66) and picnics (.53), and date (.50). Physical activities engaged in were tennis (.60), swimming (.44), and baseball (.35), but they tended *not* to participate in basketball (.32), hunting (.30), and foot-

ball (.26). It was their preference that members of the group recognize the importance of good manners (.35) and like to date (.28). The groups named were likely to be of heterosexual composition (.29) and to some extent participate in the activities of riding in cars (.29), going out of town to have a good time (.27), and walking around town (.26). They were more likely to meet at night (.31) or in the afternoon (.25).

Many of their activities were related to the school. They studied together (.54), went to school affairs (.46), met in the hallways of the high school (.45), attended school athletics (.44), went to a school sponsored recreation center (.43), and met at the public library (.42). They preferred that their group members be of the type which would not quit school (.41). The adolescents naming this type of group were likely to be in grades 11 and 12 (.32) and to have IQ's above 111 (.32) and not below 89 (.34). Also, at school they attended movies together (.39) and met outside the school building (.35).

Dominant in this factor were socioeconomic variables. Many of the activities depended upon the adolescent's being a member of a family in average or better economic circumstances. There was a definite trend for adolescents from low socioeconomic level families *not* to designate this type of informal group as their favorite.

In Factor V there are evidences of a need to satisfy heterosexual interests and a need to participate in activities away from the home and community. However, many of the activities, such as having slumber parties, listening to records, having picnics, etc., suggest that approval of the family and ability of the parents to provide certain facilities were important influences. Worth noting is the fact that much activity involved the school.

It may appear somewhat incongruous that a factor with such a high loading on having slumber parties would also have a loading on the item of naming a heterosexual group. Upon inquiry it was found that it is not uncommon for slumber parties to include boys in the early stages and the morning after.

Apparently Factor V involved a somewhat complex set of social relationships characteristic of upper socioeconomic adolescents in the last two grades of high school. Girls appeared to be most sensitive to the pattern of pressures resulting in Factor V.

It is proposed that this factor be briefly described as follows: *Pattern of pressures leading to the assumption of an upper socioeconomic, quasi-adult social role.*

Factor VI

Social conformity and a concern for appearances are central to Factor VI. Adolescents' groups influenced by this factor rejected people who ride motorcycles (.61), "park" (.52), make wisecracks (.48), desire to quit school (.38), and are noisy (.35). They tended to reject other adolescents who race cars (.29), smoke (.29) and "neck" (.25). A strong

preference was expressed for persons who are good-looking (.55), are good students (.42), work after school (.39), and are of about the same age (.31). It was important that others in the group not be afraid of being different (.30). There was a tendency for them to like people who stick up for their family (.28), nearly always agree (.27), and dress neatly (.25).

Adolescents in groups scoring high on Factor VI went places together mostly associated with the school (.48, .29, .27) and a small dairy bar (.27). They participated in only two activities, volleyball (.39) and tennis (.26).

A concern for good appearance and the rejection of a noisy, show-off type of social role seems to typify Factor VI. Undoubtedly, pressures from the school and the home have much to do with the type of social conformity expressed by this factor. At first glance, the fact that a positive value was placed on "not afraid to be different" seems a contradiction. However, it is not uncommon for adults trying to produce conformity in adolescents to point out that if one wishes to amount to anything he must not be afraid of being different from other people.

It is therefore proposed that this factor be characterized as: *Pressures toward social conformity manifested by a concern for good appearances and rejection of a noisy, "show-off" type behavior.*

Factor VII

Groups scoring high on Factor VII frequented a water gym (.66) with picnicking facilities located about three miles outside the city and the city swimming pool (.26). They also liked to go to a partially abandoned amusement park (.61) at which a miniature golf course was still in operation. Restaurants (.54) were favorite places for congregating and they liked to sit and talk in them (.28). A dairy bar (.30) favored by adolescents scoring high on Factors I and III, but to some extent rejected by those scoring high on Factor V, was mildly popular. Groups of this type were most likely to be named by boys (.58).

As groups they participated in the sports of golf (.58), football (.57), basketball (.43) swimming (.34), tennis (.32), and wrestling (.25). They went fishing (.52), hiking (.47), and hunting (.44) together. They sometimes played pool (.47), but tended to reject the activity of making candy (.26).

There was some evidence that groups scoring high on this factor may have been influenced to some extent by the school. They preferred that their group members be popular in school (.60). A recreation center (.53) housed in a school was a popular meeting place and they sometimes met outside the high school building (.41). It is not clear whether they played together at the high school stadium (.43) or went there to view athletic events.

They liked members of their group to be good looking (.37), and to work after school (.30), but they did not care if they swore (.37), would fight (.29), or got in trouble with teachers (.27). Members were acceptable

who were always "kidding" (.28), who "necked" (.26), and sometimes cheated (.26). It was not necessary that they dress neatly (.26).

The fact that groups influenced by Factor VII were most frequently named by boys and many of the activities engaged in by the groups are usually considered to be primarily of interest to males suggests that this factor may be largely the result of pressures involved in playing a masculine role. Many of the activities required that the adolescent have the freedom to move about and freedom from adult supervision frequently denied girls in our culture.

It is interesting that the groups named by adolescents scoring high on this factor were not necessarily unisexual in their composition. As a matter of fact, there was a 50-50 chance that the groups named were made up of both boys and girls. This suggests that some females may have been involved in this factor who were perhaps having difficulty in establishing their roles as feminine members of society. Worth noting also is the fact that no socioeconomic variables were related significantly to Factor VII. If this factor is an expression of the male role, apparently socioeconomic status has much less meaning for boys than it does for girls.

Factor VII gave some evidence of being positively related to the school, but some activities and values deemed quite important by the adolescents influenced by Factor II had negative loadings on Factor VII. Apparently, the adolescent male is allowed to hold some values contrary to those held by the school and upper socioeconomic families. Or perhaps, this is a manifestation of his rebellion against the restriction of such a moral code.

It is proposed that Factor VII be considered: *Pattern of pressures and needs involved in playing a masculine role.*

Factor VIII

Some kind of conformity seemed to be important in this factor. It was apparently necessary that members of the group strive to please parents and teachers. They preferred that their friends be good students (.43), plan on going to college (.36), and stay out of trouble with their teachers (.38). It was important that group members mind their parents (.40) and be in early at night (.32).

What other people thought of them appeared to be important to young people scoring high in this category. They liked people who had a good reputation (.42) and believed that good manners are important (.32). A preference was expressed for people who were not afraid to be different (.43). Apparently they were willing to maintain these standards at the expense of sometimes being considered a little "queer." Or perhaps, adults in seeking to control them explained that individuals who wish to amount to anything can not be concerned about appearing to be different from the general run of people.

They avoided the slate pit (.34) and the old county fair grounds (.35) which was a large wood lot at the edge of town. It was unlikely that they

would make candy (.31) when they got together. Their groups met at school movies (.35) and outside the high school building (.30), but to some extent avoided the school cafeteria (.26). Adolescents most frequently naming this type of group were in the ninth and tenth grades (.34).

It is somewhat difficult to interpret Factor VIII; however, it seems fairly apparent that adolescents under the influence of this factor must have been convinced of the importance of pleasing their parents and certain teachers. On the other hand, one can hardly resist conjecturing what the effect of living by such a code might have had on peer relationships, particularly at the ninth and tenth grade level. Could the adolescents influenced by this factor have been isolates because their standards made it difficult for them to establish satisfactory peer relationships? Only further, more definitive research can answer this question. In the meantime, it is tentatively suggested that Factor VIII be considered a: *Pattern of pressures resulting from adult domination and lack of emancipation from the home.*

Factor IX

Good appearances seemed very important in Factor IX. A strong preference was expressed for people who dress neatly (.68), have a good reputation (.62) and think manners to be important (.56). Frequent church attendance (.33) was desirable. Persons were likely to be rejected from their groups who told off-color jokes (.57), swore (.49), made wisecracks (.32), would fight (.49), or would gamble (.47). They preferred that members of their group be white (.66) and were likely to reject colored young people (.56). A high value was placed on minding parents (.61), and there was a tendency to prefer people who got in early at night (.27). It was important to avoid appearing as if one always had to win (.59) presumably in various kinds of games or discussions. Racing cars (.40) was frowned on.

Factor IX almost took on the appearance of being an expression of a feminine role. Perhaps it is an expression of a particular type of female role. Many of the activities important in Factor VII which appeared to be dominated by masculine interests were rejected. Adolescents of the present factor did not wrestle (.63) or play basketball (.61) and football (.55). Hunting (.53) and fishing (.45) were not at all popular. Playing pool (.53) was frowned on. Baseball (.29) and roller-skating (.27) were the only physical activities in which they engaged to any extent. They liked to go on picnics (.26).

Young people naming groups rated high on this factor came from families receiving a composite socioeconomic rating in the middle group (.40) for the community. Their fathers tended to be employed in occupations falling in the middle (.26) of the classification employed in the present study. Seventh graders (.37) were not likely to be found among those naming groups high in Factor IX.

Apparently the questionnaire failed to include any places that groups of this kind liked to go; however, they appeared to avoid some which were

listed. The playground of an elementary school (.49) located in the better part of town, school movies (.41), and the cafeteria at high school (.26) were not acceptable meeting places. The slate pit (.48) and the old fair grounds (.33) were not approved places to be seen. A dairy bar (.26) popular with adolescents of Factors I and III and avoided by those of V was to some extent avoided by young people scoring high on Factor IX.

The moving force behind Factor IX appeared to be a need for approval and status. Apparently this need found its expression in maintaining the very best of appearances, having good intentions, associating with the right people and avoiding activities which might mark one as an uncouth or perhaps unladylike person. Since groups scoring high on the present factor were most frequently named by middle-class girls, it is perhaps to some extent the manifestation of a middle class, feminine role. Apparently it does not become an issue until after the seventh grade.

Factor IX may suggest that the most frequent means for the middle class girl to climb the socioeconomic ladder is through a program of appearing to be highly virtuous, having the very best of intentions, and avoiding any activity which would detract from her femininity. On the other hand, such a pattern of activities, attitudes, and values might be an attempt to maintain status by appearing to be as different as possible from those in a lower socioeconomic classification.

It is significant to note that the school does not play a role to any extent in this factor, in spite of the fact that many of the values related positively to this factor were important in Factor II which was so closely allied to the school. There was, however, some indication that the home plays a part in the formation of the factor.

It is proposed that Factor IX be characterized as a: *Need for approval and status growing out of pressures applied by the middle class family.*

Factor X

After a thorough consideration of the variables which emerged in Factor X the meaning was so unclear that it did not seem wise to attempt an interpretation. It was the last and weakest factor to emerge from the analysis.

Discussion and Conclusions

Many of the variables in the present study had loadings on more than one factor. While an attempt is made in multiple factor analysis by means of rotation to maximize loadings on one factor and minimize them on all other factors, there are limits imposed by the very nature of the data under consideration. This is true because some activities and attitudes have varied meanings to adolescents and are used by them in different ways. As an example, dancing, with a loading on Factor I, appeared to be a means of assuming the adult role and to a considerable degree a means of satis-

fying heterosexual interests. In Factors III and V it appeared to take on the meaning of an activity by means of which relatively uncomplicated social needs are satisfied. Making candy in Factor III appeared to be an activity which is done around home and which in all probability requires the cooperation of parents. In Factor IV it appeared to be one of the few activities which can be done at the home of adolescents who are of very low socioeconomic status, while in Factors III and V making candy was one of many activities characteristic of upper socioeconomic social groups. In Factors III, V, and VII, making candy appeared to be a more feminine than masculine activity. In Factor VIII it took on the possible meaning of an activity which is refused children in a family exercising undue control. While the analyses just made of the two activities may be subject to a great deal of error, they are examples of the kinds of hypotheses which might be proposed on the basis of the present type of study.

There was evidence in the present study to indicate that socioeconomic status of the family is an important factor in connection with attitudes and activities of adolescents. It does not, however, stand out as a dominating factor. How much effect the family's socioeconomic status has upon the behavior of adolescents can not be determined from the present study. However, it does appear that a very important difference between the socioeconomic levels appears in the degree and quality of emancipation from adult control allowed the adolescent. Emancipation is likely to depend upon many factors. In all probability the role of an adult will be forced on those adolescents of lower socioeconomic status sooner because of the requirements of having to support one's self at an earlier age. But, aside from financial necessity, the way in which emancipation from adult control takes place is probably highly conditioned by standards and values depending upon culture, religion, philosophy, and education. It would seem that more important than socioeconomic status itself is the quality and degree of emancipation from adult control which appeared to be a major issue in factors not loaded to any degree on socioeconomic status.

Adolescent girls tended to mention groups that were typical of factors showing a considerable amount of association with socioeconomic status. This fact may indicate that girls tend to make distinctions along socioeconomic lines more often than do boys. Or perhaps it may be accounted for by the fact that girls tend to be under more strict adult supervision. Closer supervision and closer home ties may produce in girls a greater consciousness of status and values held by the family than is found in boys.

Noticeable is a tendency for the upper socioeconomic factors to be associated with activities centered in the school. This suggests that the school tends to encourage attitudes and standards that are characteristically upper socioeconomic in nature. "Upper socioeconomic" in this small Ohio city probably means what one would ordinarily speak of as "middle class" for the United States as a whole. In this city there was a very small "upper" class insofar as one existed at all. In the socioeconomic typing there were

no houses that could be rated "1," using Warner's method. It was most difficult to make clear-cut distinctions as to the quality of the various neighborhoods.

Chronological age had a high loading on Factor I. It had barely significant loadings on Factors VII and X, but because they were below the cutting point of $\pm.25$ they were not used in the interpretation of the two factors. The fact that chronological age was associated to any degree with only one factor out of 10 suggests that it is by no means a dominant factor in adolescent behavior and attitudes. On the other hand, grade position in school was loaded significantly on several of the factors. Of the two, then, grade position is probably more useful in predicting adolescent group behavior. This may be due to the fact that roles are assigned adolescents which are characteristic for grade levels. In other words, it is likely that social maturity is gauged by the layman and peer groups more by school grade than by chronological age.

Intelligence tended to be related to the same factors upon which socioeconomic measures were loaded, thus bearing out the considerable amount of research that shows a positive relationship between intelligence as measured by the traditional mental test and socioeconomic status. There is no evidence in the present study to throw light on the nature of the relationship between measured intelligence and socioeconomic status.

There was a tendency for factors loaded on upper socioeconomic status to be loaded also on the variable "seven or more in a group," indicating perhaps that adolescents of higher socioeconomic status tended to associate with larger peer groups than did the adolescents of lower socioeconomic status. Apparently, adolescents of high sociooeconomic level tended to make more social contacts and to be involved in more interpersonal relationships than those of a lower socioeconomic level. This fact may be of significance in the area of social skills. The upper socioeconomic status adolescent has much greater opportunity to practice the skills of social intercourse. As a result, he is likely to have a great advantage over his peers of lower socioeconomic status when it comes to effectively relating himself to others in the competitive situation which is characteristic of his social milieu.

While the naming of factors in the preceding section of the present study is in a real sense a statement of conclusions, it seems worthwhile to set forth several generalizations which grew out of the study as a whole.

1. The present study has demonstrated that in the area of psychosocial research multiple factor analysis can be a useful statistical method in the development of hypotheses and categories from empirical data all of which have had equal opportunity to influence the final outcome.

2. While socioeconomic status is significantly related to many areas of adolescent informal group attitudes and activities, it does not appear to be the dominant factor.

3. In her informal group activities and attitudes the adolescent girl is influenced to a greater degree by socioeconomic status than are boys.

4. Identical activities and attitudes may be used by adolescents to satisfy widely divergent needs.

5. Degree of emancipation from adult control appears to be a most important influence in the formation of patterns of informal adolescent group activities and attitudes.

6. The public school as a social institution tends to re-enforce upper socioeconomic values and is used to a considerably greater extent to satisfy social needs by upper-socioeconomic young people than by those of lower status.

Summary

In the present study the area of adolescence was considered through the investigation of adolescents' reports concerning activities and standards of social acceptability of informal groups with which they claimed most frequent association.

A small city in Ohio was selected in which to conduct the research project. A questionnaire was developed and administered to students in grades 7 through 12 of the public and parochial schools. In addition to data gathered in the questionnaire, intelligence quotients and socioeconomic ratings were obtained for each subject.

An adaptation of multiple factor analysis was selected as the statistical tool best suited for analyzing the data. Ten factors emerged from the analysis, nine of which were tentatively identified.

References

1. Anderson, H. H., & Anderson, Gladys L. Social development. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd Ed.) New York: Wiley, 1954. Pp. 1162-1215.
2. Jones, H. E. The environment and mental development. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd Ed.) New York: Wiley, 1954. Pp. 631-696.
3. McCandless, B. Environment and intelligence. *Amer. J. ment Def.*, 1952, 56, 674-691.
4. Warner, W. L., & Meeker, Marcia. *Social class in America: A manual of procedure for the measurement of social status*. Chicago: Science Research Associates, 1949.
5. Wherry, R. J., & Gaylord, R. H. The concept of test and item reliability in relation to factor pattern. *Psychometrika*, 1943, 8, 247-264.

GENERAL INFORMATION OF KINDERGARTEN CHILDREN: A COMPARISON WITH THE PROBST STUDY AFTER 26 YEARS

MILDRED C. TEMPLIN
Institute of Child Welfare, University of Minnesota

Many descriptive studies of child behavior were carried on in the 1920's and the 1930's. Since that time the environment of young children has changed greatly, and the results of recent but comparable studies of behavior are both similar to and different from those carried on 20 or more years ago. Since the range of children's information should be particularly sensitive to the differential impact of a changed environment, the author repeated in 1954-1955 the study carried on by Probst in 1928 with kindergarten children (2, 3).

PROCEDURE

Sample

In the present study the number, sex, age and socioeconomic distribution of the subjects in the major sample are the same as in the Probst study. Both studies included 100 children, 50 boys and 50 girls between five years, four months, and six years, chosen to be representative of the total population of Minneapolis according to their father's occupations as classified on the Minnesota Scale of Paternal Occupations. In both studies the 50 children in classes I, II, and III (professional, semi-professional and managerial, clerical and skilled trades) constituted the upper socioeconomic group, and the 50 in classes V, VI, and VII (minor clerical jobs, semi-skilled trades, slightly skilled and unskilled labor), the lower socioeconomic status group. Probst's subjects were selected from seven Minneapolis Public schools and those in this study from 11 such schools in neighborhoods appropriate to their fathers' occupations. Although Probst had Detroit Intelligence Test scores available, in the present study no such data were available and no intelligence test was given. An indirect control of intelligence is achieved in the selection of the sample.

The children in this study were questioned during their first semester in kindergarten. Since Probst questioned her subjects during the second

semester, an additional 24 kindergarten children in socioeconomic class V enrolled in the same schools as the class V subjects in the major sample were interviewed the following May.

Questionnaire

The Probst questionnaire samples 11 categories of knowledge: I. Local points of interest; II. Time and number; III. Current topics and history; IV. Natural phenomena; V. Literature and music; VI. Animals, birds and insects; VII. Plants and flowers; VIII. Occupations and industries; IX. Household arts; X. Simple mechanics; XI. Games and amusements. Each form, A and B, was approximately equal in difficulty and contained 66 questions, six in each category of knowledge.

All the Probst questions were used in the present study so that specific comparisons are possible.[1] However, 21 new questions were added: 16 replacing questions judged out-dated in categories III, IX, X, and XI, and five parallel questions sampling current information in category III. These new questions together with the number of the Probst items they may replace appear in Table 4.

In both studies the children were questioned individually: by Probst in two separate sessions within a seven-day period, in this study in a single session. Probst reported a mean of 17 minutes for the administration of Form A and of 19 minutes for Form B, while the mean time for the single interview was approximately 30 minutes.

Probst's rules and principles were followed in scoring the original and the new questions. Some specific responses not credited in the earlier study might be in this one. For example, in response to "What is the Great Northern?" not only references to the train, trains, or the depot but also references to the Great Northern Markets (a local chain) were credited. The score in the Probst study was the number of correct responses. In this study the score on the original Probst questions is referred to as the "regular" score. The score achieved by replacing certain questions with the new substitute questions is designated the "substitute" score. For both regular and substitute scores the maximum value is 66 for each form and 132 for the entire questionnaire.

Analysis of Responses

Probst reports the percentage of correct responses given by boys, girls, upper and lower socioeconomic groups, and the total sample. Certain observed differences were tested for significance. In her qualitative analysis she discusses characteristics of the children's responses. In this study similar procedures are used in the analysis of the major sample and the class V subjects.

---

[1] Probst reported a .94 (corrected to .97) correlation between Forms A and B; the comparable correlation in this study is .89 (corrected to .94).

TABLE 1

MEAN REGULAR AND SUBSTITUTE* SCORES ON TOTAL TEST, FORM A AND FORM B FOR THE MAJOR SAMPLE AND CLASS V SAMPLES INTERVIEWED IN NOVEMBER AND MAY

	Major-Sample ($N = 100$)		Class V November Sample ($N = 24$)		Class V May Sample ($N = 24$)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Regular Score						
Total	58.3	16.2	56.0	14.0	58.7	15.2
Form A	28.5	8.6	27.5	6.4	29.9	8.5
Form B	30.3	8.2	30.6	5.4	29.1	7.1
*Substitute Score**						
Total	63.2	17.5	61.2	16.0	63.7	16.7
Form A	31.0	8.7	31.0	8.7	33.6	10.0
Form B	31.6	8.1	30.5	7.7	30.7	7.5

* Score determined by replacing certain Probst questions with new items.

RESULTS

Quantitative Analysis

Total Scores

In Table 1 mean regular and substitute scores on both forms and for the combined forms are presented. The mean total score reported by Probst was 71.6, *SD* 19.05, which is higher at the .01 level of confidence than either regular or substitute total score in this study ($t = 5.25$ for the regular, and 3.24 for the substitute score). This difference cannot be accounted for by poorer test performance during the latter part of the interview in the present study, because the mean scores on Forms A and B do not differ significantly, and Form B actually yields the higher score. Although Probst does not report mean scores for the separate forms, she states that the mean for Form B was 1.6 points higher than the mean for Form A. Nor can the difference between 1928 and 1954 scores be attributed to a difference in the amount of kindergarten experience of the subjects. Although the mean scores on Form A, Form B, and the combined forms tend to be higher in the sample tested in May, in no instance does the difference approach the .05 level of confidence.

The Probst sample yielded higher scores on all categories of knowledge except category X for which no difference occurred, and category VI for which the present sample received the higher score. Using the substitute scores, the differences between the studies are significant at the .01 level for categories I, II, and VIII and at the .05 level for category VII ($t = 5.24$, 3.68, 3.47, and 2.53, respectively).

Sex Differences

Table 2 shows that the boys consistently score higher than the girls, but the differences in scores of the two sexes are statistically significant only on three categories for the Probst samples. In both studies boys receive higher scores in six of the same categories, I, III, VI, VII, VIII, and X, but girls in only one, IX. When the scores on the two studies are compared, those of the girls are not significantly different, but those of the boys are significantly higher in the Probst study.

TABLE 2

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE SCORES OF BOYS AND GIRLS IN THE PROBST AND TEMPLIN STUDIES

	Boys ($N = 50$)		Girls ($N = 50$)		
	Mean	*SD*	*Mean*	*SD*	*t*
1. Probst total score	78.0	18.4	67.4	18.7	2.85†
2. Templin regular score	59.6	16.2	56.7	15.9	0.90
3. Templin substitute score* ..	65.6	16.6	60.9	17.3	1.38
t (1 and 3)	3.52†		1.80		

* Score achieved by replacing certain Probst questions with current items.
† Significant at the .01 level.

Socioeconomic Differences

As seen in Table 3, the upper socioeconomic group consistently receives scores which are significantly higher than the lower group. In both studies the higher score of the upper group appears in each category. In the Probst study only the difference for category II is not significant at the .01 level, but in the present study only the difference for category VII reaches this level. The upper Probst sample scores significantly higher than the present upper sample, but the difference between the lower groups is not statistically significant.

Specific Questions[2]

In comparing the total samples, more subjects in the present study responded correctly on 27 questions, fewer on 98 questions, and the

[2] A table giving percentages of correct responses on each question by both sexes, both socioeconomic groups, and by the total sample in the present study and by the total sample in the Probst study has been deposited as Document number 5368 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C. A copy may be secured by citing the Document number and by remitting $1.25 for photoprints, or $1.25 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

TABLE 3

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE SCORES OF THE UPPER AND LOWER SOCIOECONOMIC STATUS GROUPS IN THE PROBST AND TEMPLIN STUDIES

	Upper ($N = 50$)		Lower ($N = 50$)		
	Mean	*SD*	*Mean*	*SD*	*t*
1. Probst total score	80.0	16.5	63.2	17.7	4.90†
2. Templin regular score	64.6	14.6	51.7	14.7	4.41†
3. Templin substitute score* ..	69.6	15.9	56.9	16.7	3.87†
t (1 and 3)	3.21†		1.83		

* Score achieved by replacing certain Probst questions with current items.
† Significant at the .01 level.

same number on nine questions. On 36 questions the differences in the number of correct responses is significant at the .01 level. More Probst subjects responded correctly on 33 and the present-day subjects on three of these. Significant differences appeared on 10 of the questions for which substitute questions were constructed.

New Questions

The new questions are presented in Table 4 with the percentage of correct responses given by both sexes, both socioeconomic groups, the major and class V samples. In several instances, because of the difficulty in constructing comparable questions, more than one substitute question was tried. In determining total scores, the question to which more children responded correctly was used. For example, as a substitute for identifying Jack Dempsey more children identified Rocky Marciano than Joe Louis so the former was used.

Qualitative Analysis

The actual responses of the children are essentially similar to those reported by Probst; yet in both studies the responses reflect the environment in which they were offered. Jackie Coogan was identified as "a boy on Lawrence's sweater," and Dagwood as "a man by our house." The present age is reflected in one child's identification of the part of a plant that is underground as the "shelter." Many identical clang associations are used in both studies. Present-day children still say that butterflies make butter; that bees make beans; that a plumber plumbs or sells plums; etc. Such clang associations still are extended in explanations when children report that a carpenter cleans or lays carpets. Concreteness is also apparent. One example is in responses to "What time is it at noon?": time for lunch, time to go to or return from school, and "When Casey Jones comes on" (a reference to a noon television program).

TABLE 4

PERCENTAGE OF CORRECT RESPONSES ON SUBSTITUTE AND PARALLEL QUESTIONS BY SEX AND SOCIOECONOMIC STATUS

Form	Category	Item Replaced	Question	Major Samples: Upper (N=50)	Lower (N=50)	Girls (N=50)	Boys (N=50)	Total (N=100)	Class V: Nov. (N=24)	May (N=24)
Substitute Questions										
A	III	1	Who is Dagwood?	34	14	28	20	24	17	25
A	III	2	What is a helicopter?	50	44	32	62	47	37	58
A	III	4	Who is Rocky Marciano?	20	20	8	32	20	21	25
A	III	5	Who is Hopalong Cassidy?	94	86	90	90	90	87	100
A	III	6	Who is Dwight D. Eisenhower?	58	44	46	56	51	54	29
B	III	1	Who is Dennis the Menace?	66	56	54	68	61	54	62
B	III	4	Who is George Mikan?	8	6	4	10	14	8	8
B	III	5	Who is Harry Truman?	12	12	8	16	12	8	8
B	III	6	Who is Adlai Stevenson?	10	8	4	14	9	4	4
B	IX	2	What do we have to have to make a cake?	88	78	88	78	83	67	62
B	IX	3	What is a mangle used for?	6	0	4	2	3	0	4
B	IX	4	Where do we keep frozen foods?	98	98	98	98	98	92	96
A	X	6	What makes an iron hot?	14	6	6	14	10	0	12
A	XI	3	You know how to play "The Farmer in the Dell." Whom does the farmer choose?	68	64	80	61	71	46	79
B	XI	1	What do we use to play golf?	72	50	72	50	61	54	50
B	XI	4	What do people play Canasta with?	22	20	24	18	21	29	29
Parallel Questions										
	III		Who is Carnival Clown?	68	64	76	56	66	71	71
	III		Who is Howdy Doody?	96	96	94	98	96	92	87
	III		Who is Cedric Adams?	72	66	64	74	69	79	71
	III		Who is Arthur Godfrey?	90	80	88	82	85	83	87
	III		Who is Miss Frances?	58	36	52	42	47	46	37

The incorrect responses of children illuminate the levels of their understanding. Lack of appreciation of cause and effect is seen in the response of several children to "How is lumber made from trees?" in which several children started with a board, stuck it into the ground, tied branches on it, put leaves on the branches and ended up with trees. The imaginative explanations of children are illustrated by the idea of one child about lightning: "I think that it's a broken sun."

Only three children did not have television sets in their homes and even they had access to one. The impact of television is apparent in the high percentage of the children who correctly identified both local and national television personalities (*see* Table 4). It is also evident in the precise, detailed information offered on the program content and schedules, and on the personalities of the performers. "He's on TV," or "He's got a show," was a frequent response in identifying a person a child did not know. The children named their favorite television programs specifically and often described them with a good deal of color and enthusiasm.

In contrast, when asked about their favorite radio programs, the children often named television programs (some of which were also on radio) or general areas such as news, music or sports. Radio was not as universally present or listened to by the children. Where it was, however, it sometimes served a definite function. For example, several children indicated that they listened to "sports with daddy," and one child said "I listen to the weather so my mother knows how to dress me."

That present-day children are, at least verbally, aware of more of the world is illustrated in responses to "What large city is closest to Minneapolis?" The percentage of correct response is about the same in both studies, but only three or four responses reported by Probst mentioned places outside the state of Minnesota—California, London, Chicago, and Washington (which may also refer to a well-known street in Minneapolis). Among the incorrect responses in the present study, two cities outside of Minnesota (Fargo and La Crosse), three states (California, Michigan and Texas), and two countries (Canada and Japan) are mentioned. Although television may influence some responses, in nearly all instances, the youngsters reported some friend or relative living or having lived in the place mentioned. Adults today are more mobile and this is reflected in the responses of the children.

Summary and Discussion

A quantitative comparison of the responses of kindergarten children in 1928 and 1954-1955 indicates that on the same questions, children in the earlier period scored significantly higher. The differences in the range of information between boys and girls seem to have decreased to a greater extent than those between upper and lower socioeconomic groups.

A qualitative comparison of the actual responses of the children points up essential similarities in child thought between the two periods. In both, children use clang associations, concrete reference, and reflect many aspects of the environment in which they live. There is substantial evidence that television is a pervasive element in the lives of present-day kindergarten children.

Because present-day children gave fewer correct responses to the same questions does not necessarily mean that they are less well informed. It probably indicates that the questions were better suited to the experiences of children in 1928. In constructing the original questionnaire, the greater number of items Probst included dealt "with objects fairly common in the environment of most urban children or with those that form topics of conversation" (3, p. 83). The findings suggest that many such questions may be less pertinent for present-day children. Nevertheless, outside of the category on current topics and history there were few questions which practically none of the children answered correctly. As the total environment changes, the specific environment of the children also shifts, but this environment has changed gradually enough over the past few decades so that a substantial part of the information sampled is not irrelevant.

An important problem inadequately met by both studies is how to describe or analyze the environment in which the study is carried on. That such a problem is inherent in most descriptive studies of child behavior has been recognized (1). An identical question asked children may not mean the same merely because of the time or place it was asked. The question showing one of the greatest shifts in correct responses was "Tell me the name of a Minneapolis newspaper." In 1928 85 per cent gave a correct name, but in 1954-1955 only three children attempted a name, and none knew one. Since most of the children reported newspapers in their homes, the decline is not evidence of unfamiliarity with newspapers. However, in 1928 Minneapolis had two major morning and two major evening newspapers. To indentify one, it was necessary to name it. In 1954-1955 the city had one morning and one evening newspaper, so the specific name was less important. In both studies, although the number of "correct" responses differed, it is likely that the children referred to the paper as they heard it referred to about them.

The over-all increased stimulation in the environment may account for the differences in some responses. For example, significantly less information about what was made in the Ford plant was given in 1954-1955. The Ford does not hold the same unique place among automobiles as it did in 1928, soon after the introduction of the new and much-discussed Model A. The St. Paul Ford assembly plant as part of the area's industry does not hold the same unique position that it held when it was three years old and an early example of Ford decentralization in assembling cars. While children of kindergarten age are acquiring much specific information, they are less able to generalize from this information than they will be when

they are a few years older. Fewer children had this particular bit of specific information which was less unique in their environment than it was in 1928.

Change in environmental context was apparent when the substitute questions were being constructed. In the category of current topics and history it was simple to substitute persons in comparable positions for president, etc., although this was no assurance of the comparability of the questions. However, in 1954 there was no popular hero with the appeal of Lindbergh; there was no child actor comparable to Jackie Coogan; and neither Rocky Marciano nor Joe Louis was as well known as Jack Dempsey had been.

The results seem to indicate a greater similarity in the environmental stimulation of present-day children. The decreased differences between the responses of boys and girls are in agreement with other recent findings (4, 5), and may reflect a disappearing double standard in dealing with children. Although still apparent in the present study, significant differences in the performance of children in the upper and lower socioeconomic groups were not as frequent. The number of subjects selected in each occupational class was identical with the Probst sample, although there has been substantial occupational shift and some upgrading since 1928. Had the 1954 occupational distribution of the city been used in selecting the present sample, differences between the performance of socioeconomic groups may have been less apparent. This greater similarity of the environmental stimulation is also suggested by the results of the comparison of the sex and socioeconomic groups in the two studies: insignificant differences were found between the groups obtaining lower scores in the earlier study (girls and lower socioeconomic children), but significant differences were found between the groups obtaining the higher scores on the earlier study (boys and upper socioeconomic children).

The responses of the children in the two studies differ chiefly in the frequency of correct responses. This may merely reflect some shift in the environment of the children. While this series of informational questions gives some indication of how at a given time and place the information of a child "compares with older children his age" (3, p. 83), it cannot necessarily be used to compare children at different places or at different times. To sample adequately the range of information of present-day children an extensive set of current questions must be devised, and, of course, revised continuously or periodically to take into account environmental change.

There is, however, substantial agreement in the qualitative aspects of the responses offered in both studies. The actual responses of the children, frequently the erroneous ones, indicate similarity in performance between the two periods. Had the questions asked in 1928 permitted more comparison of concepts, it may be that this similarity would have extended to more stability in the percentages of acceptable responses as well. The

conceptual level attained by children as well as the specific information acquired should be considered in further comparisons over time of knowledge of children. Being able to say the number of pennies in a nickel or a dime is probably at least as much related to specific experiences with money as to having a concept of the quantity "five" or "ten."

References

1. Barker, R. G., & Wright, H. F. *Midwest and its children.* Evanston, Ill.: Row, Peterson, 1954.
2. Probst, Cathryn A. Extent and range of information possessed by kindergarten A children in the Minneapolis Public Schools. Unpublished master's thesis, Univer. of Minnesota, 1928.
3. Probst, Cathryn A. A general information test for kindergarten children. *Child Develpm.,* 1931, 2, 81-95.
4. Templin, Mildred C. *Certain language skills in children: Their development and interrelations.* Minneapolis: Univer. of Minnesota Press, 1957.
5. Templin, Mildred C. Play activities of children and adolescents: I. A comparison of urban subjects in 1924, 1950 and 1955. (In preparation)

ADULT REACTIONS TO PREFERRED CHILDHOOD STORIES[1]

MARY J. COLLIER *and* EUGENE L. GAIER
Louisiana State University

Psychologists have tended to neglect the field of children's literature as an area of interest to themselves or, indeed, even to children. Observation, planned or casual, suggests, however, that some stories are so appealing to children that they warrant attention from psychologists on these grounds alone. Particularly striking is a child's intense and persistent preference for his favorite story as if its repetition had important personal meaning to him. One wonders what that meaning is and how one might investigate it.

The psychological literature is of limited help in this connection. Freud (4, 5), Fromm (6), Jones (7), Rank (10), and others have dealt with well-known folk tales and myths and the possibly universal meaning in their recurrent themes and symbols. However, their treatment is chiefly theoretical and without reference to data from people who like the stories, to individual differences in taste, or to circumstances conducive to preference.

Social anthropologists like Benedict (3) and Malinowski (8) deplore the tendency of these writers to interpret myths and folk tales in universal terms. They argue that themes and symbols, apparently similar, take on different meanings in different cultures and that these meanings are never universal. They do claim, however, that regardless of cultural context these stories have the invariable function of codifying and conserving the morals and traditions peculiar to a given society. But, like the psychoanalysts, they indicate little concern for individual reactions.

Accenting literary rather than specifically psychological considerations, Arbuthnot (1, 2) and Meigs *et al.* (9) have discussed the nature and

[1] The writers wish to express their appreciation to Dr. Irwin A. Berg for his enthusiastic support of this project, and to Mr. William D. Leipold for his assistance.

function of certain children's stories. They refer also to publishers' and librarians' records as indices of story popularity but remark upon their limited value in establishing the true preferences of children themselves.

The writers are interested in studying the somewhat neglected individual reactions of children to their favorite stories. Therefore, following preliminary work in a nursery school setting, a series of studies was planned for that purpose. Attention was focused on two central questions, "What stories do children like?" and "Why do they seem to like them?" Various approaches, types of subject, and kinds of analysis were planned throughout the series. The present study is the first of the series to be reported.

Subjects and Procedure

A total of 264 students, 80 men and 184 women, in the Educational Psychology sequence at Louisiana State University were asked to summarize their favorite childhood story, to report the circumstances of first encountering it, and to compare past and present impressions of it.

Table 1

CLASSIFICATION OF ALL STORIES

Classification	Women		Men	
	N	%	*N*	%
Fairy tales	109	59.9	24	30.0
Fiction	49	26.6	32	40.0
Animal stories	16	8.0	15	18.8
Religious stories	7	3.8	8	10.0
Poetry	2	1.1		
Biography	1	.6	1	1.2
Total	184	100.0	80	100.0

Results and Discussion

The 264 stories obtained were grouped according to Arbuthnot's (1) classification system. They fell into the categories of fairy tales, fiction, animal stories, religious stories, poetry, and biography, in that order of preference (Table 1). Since the overwhelming majority of the subjects (women, 86.4 per cent; men, 70 per cent) chose stories in the first two categories, only these will be discussed here.

The fairy tale category included both the old, anonymously-composed folk tales, fables, myths, and epics, like *Little Red Riding Hood, The Grasshopper and the Cricket, Jason and the Golden Fleece,* and the so-called

"new tales of magic" like *The Wizard of Oz, B'rer Rabbit,* and *Little Black Sambo.* More than half of the women's preferences were fairy tales with *Cinderella* (chosen by 32) and *Snow White and the Seven Dwarfs* (chosen by 17) clearly the most popular. Less than a third of the men named fairy tales. There was no one fairy story as popular with men as with women. The male selections also showed a contrasting taste for stories in which the chief characters were male or, sometimes, animal, as in *Jack and the Beanstalk* and *The Three Little Pigs.*

Most choices classified as fiction were realistic stories about what Arbuthnot (1) calls the "here and now," that is, "about contemporary life in our own country." Among the men the most frequently mentioned stories were *Tom Sawyer* and *Robinson Crusoe* (seven and five choices, respectively). Most popular among the women were *Heidi,* sometimes called a "modern Cinderella story," *Little Women,* and *The Bobbsey Twins* (eight, four, and four choices, respectively).

The themes, major characters, and story resolutions in the fairy tale and fiction groups will now be considered. Regarding the former, as might be expected, the fairy tale characters, generally, were less realistically portrayed than those in fiction, and the fairy tale resolutions tended to be through magical means. However, there were sex differences in specific choices.

In the fairy tales most often chosen by the women, the heroines like Cinderella and Snow White tended to be naive, helpless, and persecuted by some malevolent, older woman like a wicked stepmother, witch, or bad fairy. Not through their own actions, but through those of unusual agents like a fairy godmother, a good fairy, or dwarfs, they were protected from their persecutors. Finally, because of their beauty, virtue, or suffering, they were rescued for all time by a prince of somewhat dim personality and with no other apparent function.

The selections of the men, who chose fairy tales about half as often as the women, usually had dominant male or animal characters who were threatened by giants or by impersonal physical dangers, rather than by older women. In fact, when an older woman was present, she was likely to be benign or helpful as in *Jack and the Beanstalk* where the giant's wife protects Jack. Their resolutions were less passive and less magical than those in the women's stories, less, too, a function of virtue, physical attraction, or suffering. Rather, the outcomes were achieved through physical daring, strength, or ingenuity, as in *The Gingerbread Man, Pinnochio, Jason and the Golden Fleece, Robinhood, The Story of Roland,* and *Hansel and Gretel.*

A consideration of the fiction choices shows that the characters, themes, and resolutions here were generally more realistic than those in the fairy tales. Sex differences, however, persisted.

The most salient feature of the men's stories was the absence of any leading female characters but a pronounced emphasis on one male character

and on themes of adventure, of self-assertion, and of overcoming physical obstacles. The chief character in the men's stories tended to be an adult male, alone or in competition with other men, as in stories like *Robinson Crusoe.*

In contrast to the fiction stories chosen by the men, those of the women usually dealt with leading characters who were children, rather than adults; who might be multiple or single; and who might be, unlike the male pattern, either boys or girls. The women's themes dealt chiefly with human relationships, as such, in stories like *Little Women, Heidi,* or *Billy and His Aunt Minnerva;* with the mastery of impulses and feelings and with moral values, as in *Oliver Twist* and *The Break in the Dyke.* Themes of adventure, self-assertion or mastery of the environment were infrequent. There was less activity, less solitude, less independence, and less reliance on personal resources in the women's stories but more humor and more emphasis on human limitations, philosophically accepted.

For women, story choice was significantly related to these variables, at the 1 per cent level of confidence: first, that the story was heard at, or before, the age of 6 years; second, that the story had colored pictures; third, that the story had a happy ending (Table 2).

TABLE 2

VARIABLES ASSOCIATED WITH STORY PREFERENCE

	Tetrachoric Coefficients	
Variable	*Male* $N=80$	*Female* $N=184$
Happy ending	.29	.41**
Pictures	.26	.24*
Colored pictures	.25	.40**
Heard at or before the age of 6	.24	.38**
Heard between the ages of 7 and 10	.40**	.20

* Significant at the 5 per cent level.
** Significant at the 1 per cent level.

For men, none of the preceding variables was significantly related to story choice; encountering the story between the ages of 7 and 10 years was, however, a highly significant, selective factor.

Several factors might account for this sex difference in the age level at which the preferred story was first heard. First, the women, with earlier language development, might have heard, or read, stories at an earlier age; then, being verbally facile, they might have formulated and remembered them better than men. Second, because of the cultural emphasis on feminine passivity-dependence and on masculine activity-independence,

or because of some biological tendency in this direction, the dependent, childish stage might have been more acceptable and therefore more pleasantly and completely recalled for women. Third, because of a greater cultural emphasis on "realism" for men than for women, the men might have had a greater tendency to forget of conceal any interest in the stories of open make-believe. Fourth, the so-called "Oedipal" and "latency" periods usually correspond to the 3 to 6 year and the 7 to 10 year spans, respectively. These were preferred story periods for women and men, respectively, suggesting that the women might have been more accepting of, or were still more involved in, conflicts stemming from the Oedipal period and the men, those of latency.

That the women should prefer stories with happy endings more frequently than the men might be related to these factors. First, they chose stories from the younger age period where stories may simply have happy endings more frequently than do those for older children. Then, also, stories with happy endings were perhaps read, or given, to girls more often than to boys for "protective" reasons. Perhaps, also, regardless of their story repertoire, girls may have simply preferred happy fantasy more than boys either because it was a more congenial, or a more approved, method of mastering their particular anxieties. The more active methods were possibly more congenial to, or culturally more approved for, the boys.

That women should prefer stories with illustrations, especially those that are colored, might be explained by the fact that stories for younger children, most often chosen by the women, simply are more copiously and colorfully illustrated than those for the later age groups. Another possibility is that cultural patterns may differentially reinforce men and women for attention to decorative detail, or a native sex difference of this kind may be operating. Perhaps, too, the preferred story, through "halo effect," may be endowed retrospectively with colorful illustrations by women more often than by men because of a differential flair for, or value placed upon, them.

Negative correlations (significant at less than the 1 per cent level) were found between age at which the story was first encountered and the age for which it was recommended for both men and women (men, —.82; women, —.81; total group, —.83). Thus, the preferred story was usually recommended as appropriate for a child older than the subject when he first encountered it. Speculations about this relationship follow. Perhaps the subjects were, or saw themselves as, precocious; hence, they adopted the attitude, "Other children could not appreciate my favorite story as early as I did." It is also possible that they may have made some consistent initial error in their age-placement of the story and then made a consistent effort to revise the error in prescribing an appropriate age. Perhaps, too, as children, the subjects might have selected a favorite story which initially was more appropriate for an older child and then made a corresponding allowance in retrospect. Finally, the story may have been disturbing,

though preferred, its disturbing quality later attributed to a premature encounter with the story and the recommendation then made that it be used for older children.

When present and past attitudes toward the preferred story were compared, they appeared generally consistent with one another since about two-thirds of both men and women considered their old favorite to be their present preferred story for children (Table 3). A few had changed their minds and the remainder were uncertain about their preference for the story now. Possible explanations for this finding are as follows: First, past and present attitudes might have been consistent, as reported, or the report might have reflected, rather, some effort to maintain the appearance of consistency. Second, although the stories were preferred in the present, as in the past, adult reasons for liking them might have differed from those of childhood. Third, previous and past judgments may simply have contaminated one another.

TABLE 3

RESPONSES TO THE QUESTION: IS THIS STILL THE CHILDHOOD STORY YOU LIKE BEST?

	MALE		FEMALE		TOTAL	
Response	*N*	%	*N*	%	*N*	%
Yes	55	68.8	121	65.8	176	66.7
No	3	3.7	10	5.4	13	4.9
Uncertain	22	27.5	53	28.8	75	28.4
Totals	80	100.0	184	100.0	264	100.0

The possibility was indirectly explored that other people who liked the story exerted an influence on these subjects. However, about one-third of the women and almost one-half of the men indicated that the stories they had preferred as children had not, then, been the favorites of anyone else. Even so, almost half of both sex groups reported that their preferred stories had also been the favorites of siblings and peers. Only a few subjects mentioned, as significant, the attitudes entertained by older persons toward their favorite stories; when they did so, the older persons were almost invariably women. The results have certain implications. Men were, or recalled that they were, more independent of others' attitudes in selecting their favorite story in childhood. The preferences of siblings and peers carried more apparent weight in determining our subjects' choices than did those of older persons. Finally, when the judgments of older persons did affect the subjects' childhood choices, they were likely to be those of an older woman rather than of an older man.

Summary and Conclusions

In summary, this study is one of a series planned to explore the responses of children to their favorite story at several stages of development. Its purpose was to study the characteristics of stories preferred in childhood by college students and evaluated in retrospect.

The subjects, 264 college students, were asked to recall and summarize their favorite childhood stories and to answer questions about past and present impressions of them. Responses were analyzed for differences in age at which the story was encountered, major characters, persons introducing the stories, major themes, and differential treatment of similar themes.

Results show that women chose stories stemming from the Oedipal period, involving evil mother figures, benign but active males, and persecuted, passive young women. Their stories had been shared more with loved persons, covered a restricted number of themes, and had happy, often magical, endings. Men preferred stories encountered independently during the latency period with themes of adventure, problem solution, and self-assertion. Story endings and evaluations of the stories were more obviously reality-oriented in the male group; women figures, though rare, were almost exclusively kind and maternal.

In conclusion, types of story preferred and the manner of reporting them appear to reflect cultural, and possibly biological, sex-role expectancies.

References

1. Arbuthnot, May H. *Children and books.* New York: Scott, Foresman, 1947.
2. Arbuthnot, May H. *Time for fairy tales.* New York: Scott, Foresman, 1952.
3. Benedict, Ruth. Folklore. In *Encyclopedia of the social sciences.* New York: Macmillan, 1937. Vol. 5.
4. Freud, S. *Introduction to psychoanalysis.* New York: Boni & Liveright, 1920.
5. Freud, S. *The interpretation of dreams.* London: Hogarth, 1953.
6. Fromm, E. *The forgotten language.* New York: Rinehart, 1951.
7. Jones, E. Psychoanalysis and folklore. In E. Jones (Ed.), *Essays in applied psychoanalysis.* Vol. II. London: Hogarth, 1951.
8. Malinowski, B. *Myth in primitive psychology.* New York: W. W. Norton, 1926.
9. Meigs, C., Eaton, A., Nesbitt, E., & Viguers, R. *A critical history of children's literature.* New York: Macmillan, 1953.
10. Rank, O. The myth of the birth of the hero. *J. nerv. ment. Dis.*, 1914. (Monogr. Series No. 18)

A TEST ANXIETY SCALE FOR CHILDREN[1]

SEYMOUR B. SARASON, KENNETH DAVIDSON, FREDERICK LIGHTHALL *and* RICHARD WAITE
Yale University

For a number of years studies were done at Yale by Sarason and others (1, 2, 4, 5, 6, 7, 10, 11, 12, 13) on the effects and correlates of test anxiety. In those studies subjects were college students. For a variety of reasons it was decided to study this anxiety in school age children. Among the major reasons were (a) the unduly selected nature of a college population, limiting whatever generalizations one can make, (b) the practical difficulties one faces in securing the continual cooperation of such subjects, (c) the opinion that test anxiety was a reaction which long antedated entrance to college, (d) indications in the literature (8; 9, p. 168) that fear of school failure was one of the most common of worries or fears among children, and (e) the hope that the valid discernment of this anxiety would have more than a little practical and clinical value in the school setting. The present report concerns the initial findings of a long term project on the measurement and correlates of anxiety in children of school age. It is hoped that other research workers may find the test anxiety scale a helpful tool.

PROCEDURE

The initial form of the childrens' test anxiety scale consisted of 43 items concerned with attitudes toward and experiences in test and test-like situations.[2] As in the case of the college scale, but never made explicit in

[1] The project of which this study was a part is being supported by a grant from U.S.P.H.S. We wish to thank Mr. Joseph Foran, Superintendent of Schools in Milford, and Dr. Elmer Hagman, Director of Research and Guidance in Greenwich, for the support which they gave to the project. We owe a particular debt of gratitude to the teachers and principals of the schools used for the time and effort they gave in the collection of data. The teachers contributed a good deal of their own time in making ratings, for which we are indeed grateful. We wish also to thank Dr. John Doris for his assistance. At all stages of this project Dr. Robert Abelson was of great held to us in statistical matters.

[2] Copies of the shortened form of the scale for research purposes are available upon request. This revised form consists of 30 of the items of the original scale. In subsequent studies with the shortened scale the same results as those reported in the present article were found. We chose to present the data on the longer scale because a much larger number of cases were used and give one a better picture of the stability of the findings.

the published articles, the choice of items was guided by consideration of Freud's (3) three criteria of the anxious reaction: it is unpleasant, there are physiological concomitants, and it involves conscious awareness—anxiety is a conscious experience. Below are several of the items:

> When the teacher says she is going to find out how much you have learned, does your heart begin to beat faster? (Hand on heart—illustrate).
>
> When the teacher wants to find out how much you have learned, do you get a funny feeling in your stomach? (Hand on stomach—illustrate).
>
> When the teacher asks you to write on the blackboard in front of the class, do you sometimes feel that the hand you write with shakes a little?
>
> Do you worry a lot *before* you take a test?
>
> Do you worry a lot *while* you are taking a test?
>
> After you have taken a test, do you worry because you think you did not do well?
>
> When the teacher says she is going to give the class a test, do you become afraid that you are going to fail—that you will do poorly?
>
> When the teacher says she is going to give the class a test, do you get a nervous (or funny) feeling?
>
> While you are taking a test, do you usually think you are not doing well?

The scale was group administered. Each item was read to the class and the children were instructed to encircle a "yes" or a "no" on answer sheets. Preceding administration of the scale the following was said: "We are not giving you a test. When the teacher gives you a test, there are right and wrong answers. We are going to ask you questions about how you feel and what you think but there are no right or wrong answers. When we ask you a question, think about it and then put a circle either around yes or no. Maybe you will put a circle around yes and your neighbor puts one around no—that is because you feel differently. It does not make any difference what anybody else does. You think about each question and then put a circle around a yes or a no. Remember, there are no right or wrong answers. We want to know what you think and feel. Try to be as honest as possible." The scales were always administered by non-school personnel. The teacher of the class was never present. The children were told that no one but the person administering the scale would ever see the child's answer sheet.

In addition to the test anxiety (TA) scale each child was rated by his teacher on a 17-item scale of anxiety behavior in the classroom. Below are some of the items which were given to the teachers:

> Does the child exhibit unwarranted fidgeting (e.g., squirming, restless behavior) when called upon to recite in class?
>
> Does the child tend to become upset or anxious when a test is announced in class or he is called upon to recite?

Does the child function better when working alone than when working before the class?

Does the child work better in a situation in which he can take his time than in one in which there is time pressure?

Finally, each child was given the Otis Alpha group intelligence test as well as the Stanford Achievement test.

Before proceeding further, it seems appropriate to describe briefly certain reactions of the children to the TA scale. First, when the examiner walked into the classroom carrying the answer sheets, it almost always happened that at least one child said: "We're going to get a test!" This was even true in the earliest grades where the advisability of giving the scale had been questioned because tests were never given. Equally as striking was the fact that following administration of the TA scale, some children could be counted on to ask: "When will we find out how we did?"—this despite the instructions. These incidental observations indicate how early in life children in our culture become test-oriented and suggest how pervasive and interfering such an orientation may be. By interfering, we here mean the tendency to adopt a set which is explicitly not called for in the situation.

TABLE 1

BREAKDOWN OF TOTAL SAMPLE BY GRADE AND SEX CATEGORIES

Sex	*Grade 2*	*Grade 3*	*Grade 4*	*Grade 5*	*Totals*
Milford					
Boys	277	205	241	205	928
Girls	221	181	193	174	769
Total	498	386	434	379	1697

SUBJECTS

Children in grades 2 through 5 in six schools in nearby Milford were studied. Table 1 breaks down the subjects by grade and sex. Two schools in Greenwich ($N = 600$) were also used but, since the results are essentially the same for the two school systems, only the results for Milford will be presented. It is important to emphasize that although Milford and Greenwich differ markedly in a number of ways (e.g., social class of the population), the findings in both systems were essentially the same.

RESULTS

Reliability. Table 2 presents the findings for test-retest (two-month interval) reliability. Table 3 presents the data for split-half reliability. For both procedures a satisfactory degree of reliability is obtained.

TABLE 2

TEST ANXIETY QUESTIONNAIRE TEST-RETEST RELIABILITY

Grade	*No. of Classes*	*No. of Subjects*	*r*	SE_r*
5	3	73	.82	.040
4	3	77	.76	.049
3	2	54	.44†	.110
2	4	90	.76	.045
All grades	12	294	.71	.029

* SE_r (after Soper in Snedecor).

† One *S* obtained a score of 1 in test and a score of 41 on retest. If this *S* is omitted *r* for third grade becomes .65 and SE_r = .08.

TABLE 3

TEST ANXIETY QUESTIONNAIRE SPLIT-HALF RELIABILITY

Grade	*No. of Classes*	*No. of Subjects*	*r* (uncorrected)	*r** (corrected)
5	8	208	.788	.881
4	8	215	.795	.886
3	8	204	.816	.899
2	8	198	.695	.820

* Corrected by Spearman-Brown formula (Peatman).

Increase in TA score with grade. It was anticipated that TA scores would tend to increase with grade. This tendency was found to be statistically significant (Table 4) although it is clear that the increase in score is not a regular or steady one.

Correlation between TA and IQ. On the basis of previous studies with college students, as well as the assumption that the test anxious reaction

TABLE 4

CHI SQUARE COMPARISON OF TEST ANXIETY SCORES BY GRADE
(Milford School System, N = 1697, Mdn. Score = 12–13)

	GRADE					
	2	3	4	5	χ^2	p
No. of cases above median	213	192	198	226	26.9	< .01
No. of cases below median	285	194	236	153		

was primarily interfering in its effect, it was anticipated that the correlation between TA and IQ would be small, negative, but significant.[3] Going from grade 2 to 5 the correlations, all significant at the 1 per cent level, are −.196, −.214, −.275, and −.284. These findings are essentially the same as those reported by Zweibelson (14) with the Otis Alpha and Beta. He found, in addition, that the TA correlated significantly less with the Davis-Eells Problem-Solving games than with either the Otis Alpha or Beta—a finding predictable from the rationale of the Davis-Eells Games. Because of the high correlation between IQ and mean achievement score in the present sample, it is not surprising that the correlations between TA and mean achievement are very similar to those reported above. Going from grades 2 to 5 the correlations are: −.186, −.249, −.270, and −.294 (all significant at the 1 per cent level).

TABLE 5

CORRELATIONS OF TEST ANXIETY AND TEACHERS' RATINGS, CLASS BY CLASS, AVERAGED FOR THE VARIOUS GRADES IN BOTH SCHOOL SYSTEMS

Grade	*No. of Classes*	*Total N*	*Average r**	*Significance of r* p
Greenwich				
2	6	142	.31	$< .001$
3	7	160	.21	$< .02$
4	6	142	.29	$< .001$
5	6	139	.27	$< .01$
All grades	25	583	.27	$< .001$
Milford				
2	17	469	.10	$< .05$
3	15	377	.09	$< .10$
4	15	417	.34	$< .001$
5	14	365	.18	$< .01$
All grades	61	1628	.18	$< .001$
Total Both Schools				
All grades	86	2211	.20	$< .001$

* The average r was obtained by Fisher's method of transforming the original r for each class into z and then obtaining the weighted average value of z and reconverting to r.

Correlation of TA with teachers' ratings of anxious behavior. Although we realized the difficulties involved in utilizing teachers' ratings, it was felt that it would speak for the face validity of the TA scale if a significant rela-

[3] In the college study (11) where this finding was reported the correlation between TA and IQ was reported as +.209. A study of the scattergram given in that article will indicate that the positive correlation is a function of how the scattergram was constructed. It is clear from the scattergram that high TA goes with low IQ score.

tionship could be found between TA (a child's self-rating, so to speak) and his teacher's independent rating. Table 5 presents the data indicating that a significant relationship does obtain. It will be recalled that in the previous paragraph it was reported that the correlations between TA and IQ (and between TA and mean achievement) tended to increase as one went up the grades. The opposite occurs in the case of the correlations between teachers' ratings and IQ.

	Grade 2	*Grade 3*	*Grade 4*	*Grade 5*
TR vs. mean achievement	—.419	—.353	—.361	—.127
TR vs. IQ	—.301	—.300	—.176	—.125
TA vs. mean achievement	—.186	—.249	—.270	—.294
TA vs. IQ	—.196	—.214	—.275	—.284

In Greenwich the decrease with grade in the correlation between teachers' ratings and mean achievement is even more striking, although a similar steady decrease in the correlation with IQ is not found. It would appear, therefore, that, whereas the teachers' estimate of a child's anxious reaction decreasingly predicts a performance criterion, the children's self-estimate tends increasingly to predict the criterion. It may be that in making their ratings teachers are largely influenced by how they think children are actually performing academically and that such judgments are far more valid in the early, in contrast to the later, grades. It may also be that there are more (and more glaring) overt indicators of anxiety in the earlier grades, i.e., the anxious child in grade 2 is more easy to spot by the teachers than the anxious child in grade 5.

In the course of subsequent studies, as well as on the basis of discussions with many teachers, we have come to the opinion that teachers frequently find it difficult to believe that a bright child may also be one who sees himself as anxious about performance. If a child is considered by the teacher as an average or below average performer, the teacher seems to be much more aware of the possibility of the presence of anxiety than in the case of the clearly bright child whose classroom problem-solving behavior appears perfectly adequate to the observer. If there is in fact less chance for the anxiety of the bright child to be recognized, there would seem to be less chance for it to be reduced in strength.

TA and occupational level. The occupational level of the parent was obtained from information given by the parent when the child was first entered in school. In each case where such information was available, the parent's occupation was put in one of seven categories—1 representing the highest and 7 the lowest category. There was satisfactory agreement between two independent raters in making these categorizations. However, it was clear that the few words (or the one word) given by the parent in describing his occupation was frequently unrevealing insofar as type or

complexity of job was concerned. The average correlation (z transformation) between TA and occupational level is .119 which is significant at the .001 level. It seems reasonable to conclude that occupational level is *one* of the factors significantly related to TA. That this correlation is not larger may well be due to the inadequacies of the raw data mentioned above.

The relationship of TA with reading, arithmetic, and behavior problems. The "diagnosis" of an academic and behavior problem was made by the teacher. A chi square analysis (problem vs. no-problem cases) revealed that individuals who had either a reading, or an arithmetic, or a behavior problem obtained significantly higher TA scores than no-problem cases.

Discussion

The expected increase in TA score with grade, the negative correlation of TA with IQ, the significant correlation between TA and teachers' ratings —these and other results given above support the conclusion that test anxiety is a measurable and meaningful variable. When one considers the pervasive use of tests in our culture and the ways in which they determine the lives of the people who take them, it would be expected that the testing situation would engender an anxious reaction in more than a few individuals.

The significant and negative correlation between TA and IQ is deserving of special discussion. At first glance this finding might be interpreted simply as indicating that the less intelligent (i.e., lower test scoring) individual experiences anxiety in testing and test-like situations because it is and has been a situation in which he encounters difficulty and failure. There are several factors which indicate that such an explanation may be an oversimplification. For one thing the negative correlation was also obtained in the two Greenwich schools which were explicitly chosen because they represent an above average group intellectually. While the above explanation could still be offered for the Greenwich sample, it is clear that the term "less intelligent" could not be used in any absolute sense but in a relative way, i.e., relative to the group studied. In the previously studied college groups, which were clearly intellectually superior and where there was also a restricted range of test scores, the negative correlation also was found. In addition, whereas the high anxious college students received significantly lower aptitude test scores than the low anxious students, the former received significantly higher course grades than the latter. Zweibelson's study with a Milford sample rather strongly suggests that the size of the correlation between TA and IQ varies as a function of the type of test, mode of administration, etc.—findings in keeping with those obtained with the college groups.

Still another consideration which cautions one against simple explanations of the negative correlation between TA and IQ lies in the entirely

reasonable assumption that *the IQ score, obtained as it is in a test situation, already reflects the interfering effects of anxiety.* It is precisely because of this assumption that any explanation of the negative correlation between TA and IQ must at least consider the difference between such terms as "the less intelligent" and the "lower test scoring" individual.

There are, of course, many bright children (and college students) who have high TA scores the sources of which are not explained by any intellectual inadequacy. It is one of the main points of interest of this project to attempt to understand the sources of high anxiety in children who, from the standpoint of objective indices of performance and achievement, have experienced "success" rather than "failure."

We have here presented initial findings on a test anxiety scale for children. This scale was developed as part of a long term project concerned with the development of anxiety in children. Other studies under way concern the relationship of test anxiety to learning under varied conditions, to direct observations in the classroom, to parental attitudes and report, to behavior in the Rorschach situation, to other anxieties, and to the assignment of blame in failure situations. In addition cross-cultural comparisons are being carried out in England and Australia.

References

1. Cox, F. N., & Sarason, S. B. Test anxiety and Rorschach performance. *J. abnorm. soc. Psychol.*, 1954, 49, 371-377.
2. Doris, J., & Sarason, S. B. Test anxiety and blame assignment in a failure situation. *J. abnorm. soc. Psychol.*, 1955, 50, 335-338.
3. Freud, S. *The problem of anxiety.* New York: W. W. Norton, 1936.
4. Gordon, E. M. Self esteem, need for achievement, test anxiety and performance. Unpublished doctor's dissertation, Yale Univer., 1955.
5. Gordon, E. M., & Sarason, S. B. The relationship between "test anxiety" and "other anxieties." *J. Pers.*, 1955, 23, 317-323.
6. Mandler, G., & Sarason, S. B. A study of anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, 47, 166-173.
7. Mandler, G., & Sarason, S. B. The effect of prior experience and subjective failure on the evocation of test anxiety. *J. Pers.*, 1953, 21, 336-341.
8. Pintner, R., & Lev, J. Worries of school children. *J. genet. Psychol.*, 1940, 56, 67-76.
9. Pressey, S. L., & Robinson, F. P. *Psychology and the new education.* New York: Harper, 1944.
10. Sarason, S. B., & Gordon, E. M. The test anxiety questionnaire: scoring norms. *J. abnorm. soc. Psychol.*, 1953, 48, 447-448.

11. Sarason, S. B., & Mandler, G. Some correlates of test anxiety. *J. abnorm. soc. Psychol.*, 1952, 47, 810-817.

12. Sarason, S. B., Mandler, G., & Craighill, P. G. The effect of differential instructions on anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, 47, 561-565.

13. Wiener, G. The interaction among anxiety level, stress instructions and level of difficulty. Unpublished doctor's dissertation, Yale Univer., 1954.

14. Zweibelson, I. Test anxiety and intelligence test performance. *J. consult. Psychol.*, 1956, 20, 479-481.

EXTINCTION OF THE SMILING RESPONSE IN INFANTS AS A FUNCTION OF REINFORCEMENT SCHEDULE[1]

YVONNE BRACKBILL
Stanford University[2]

Scheduling of reinforcement has been investigated extensively in relation to the conditioning of non-social responses. The results of these studies indicate in general that a schedule by which reinforcement follows every response is less effective in maintaining performance of that response during extinction than is a schedule by which reinforcement follows only some of those responses (4).

This investigation attempts to extend the study of frequency and patterning of reinforcement to the area of social learning and to cortically immature subjects. The purpose of the research is to evaluate the relative efficacy of intermittent as opposed to regular reinforcement upon frequency of smiling in infants.

The use of an instrumental conditioning paradigm in this study is in contrast to past studies of the conditioning of the smiling response (2, 7, 8), all of which, though differing in methodology, have utilized a classical conditioning paradigm in conceptualizing results. The inappropriateness of this framework in accounting for response change over time is most clearly illustrated in the study by Dennis (2).

METHOD

Subjects

Eight normal infants between the ages of three and one-half to four and one-half months served as *Ss*. Six were males and two, females. All came from intact middle class homes located within 10 miles of Stanford University.[3]

---

[1] This report is based upon a dissertation submitted to the Department of Psychology, Stanford University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer wishes to thank Professors Albert Bandura, Quinn McNemar, Frances Orr Nitzberg, Robert Sears, William Verplanck, and C. L. Winder for their suggestions and help.

[2] Now at National Institute of Mental Health, Bethesda, Maryland.

[3] The author is greatly indebted to Drs. John Anderson and Bruce Jessup for their cooperation in referring subjects.

This particular age range was selected with the consideration in mind that the subject had to be old enough to remain awake for a short time period after feeding, yet not old enough to respond differentially to "mother" vs. "others." Other requirements for selection of *Ss* were: (a) that the infant not cry so often, with such intensity, and for so long a time that sessions very frequently had to be terminated, with consequent lack of experimental progress; (b) that the infant show an operant rate of at least two responses per five-minute interval; and (c) that the infant be able to maintain a supine position for intervals of five minutes without persistent struggling to regain the prone position.

Experimental Procedure

Ss were assigned to the regularly reinforced (RR) group or to the intermittently reinforced (IR) group in consecutive order of their acceptance as subjects. The first *S* was assigned to the IR group; the second, to the RR group; and so on.

For both groups, the experimental procedure was divided into three periods: In the first or *operant* period *E* stood motionless and expressionless at a distance of approximately 15 inches above *S*, and observed him for eight five-minute intervals to ascertain the operant level of smiling. In the second or *conditioning* period, reinforcement was meted out, contingent upon *S*'s smile. Specifically, as soon as *S* smiled,[4] *E* smiled in return, began to speak softly to *S*, and picked it up. *S* was then held, jostled, patted, and talked to for 30 seconds before being replaced in the crib.[5] In the third or *extinction* period, the procedure was again observation without reinforcement, as in the operant period. During extinction, *Ss* were observed for 13 five-minute intervals.

During all three periods, the basic interval for determining rate of responding was five minutes, and in the discussion to follow, *five-minute interval* (or simply, *interval*) refers to one five-minute period of continuous experimentation. The term *rate of response* refers to number of responses per interval. Also, the terms *response* and *trial* will be used interchangeably.

The two groups of *Ss* differed in respect to the reinforcement schedules used within the conditioning period. The RR group was maintained on a regular reinforcement schedule during the entire conditioning period. The IR group was maintained on a regular reinforcement schedule until each *S* had responded at maximum rate (*see below*) for 10 consecutive intervals. Immediately after the tenth interval at maximum rate, *Ss* were switched to a 2:1 randomized variable ratio reinforcement schedule for

[4] Interjudge reliability concerning the decision as to whether a smile had or had not occurred was found to be .975, using a total *N* of 970 such decisions in the formula: twice number of agreements / total number of judgments. This reliability check was done prior to the experiment, with the assistance of Mr. Thomas Milburn.

[5] The reinforcement procedure was recorded on 12 feet of 8 mm. film.

a total of 60 responses (hence, 30 reinforcements), then to a 3:1 schedule for a total of 45 responses, and finally to a 4:1 schedule for 20 responses. Separate randomizations were used for each subject.

Maximum rate of response was determined as follows. Forty-five seconds were required for the administration of each reinforcement plus its accompanying events. (Five seconds were required for picking *S* up; 30 seconds for reinforcement; five seconds for putting *S* down; and five seconds for recording.) Therefore, no more than six responses could occur and be reinforced during any five-minute interval. *Maximum rate of response* was defined as no fewer than four responses per interval. That particular trial or response that marked the beginning of maximum response rate will be referred to as *criterion*.

Because of intersubject variation in total number of responses to criterion and because it was desired that the experimental results be a function only of reinforcement schedule and not a function of reinforcement schedule plus number of emitted responses, it was necessary to match the two groups for total number of trials during the conditioning period. The method of matching—like the method of group assignment mentioned earlier—was an individual one and was done in consecutive order of admission to the experiment. Specifically, *S* No. 2 (RR group) was matched for total number of trials with *S* No. 1 (IR group); *S* No. 4 was matched with *S* No. 3, and so on.

General Procedure

Each infant was placed on a schedule of social deprivation during the entire experimental period. Social deprivation was defined as minimized social and body contact between the infant and any other person in its environment, excluding *E*. In effect, this meant elimination of all social and body contacts that were not absolutely necessary for *S*'s well being.

Experimental conditions were kept as standard as possible for all sessions and for all *S*s. *E* always wore a white laboratory coat. *S* was placed in a supine position near the open edge of his crib; placement of *S* and the crib was the same for all sessions. Source of light was standardized, and variance in intensity minimized. At the beginning of each session, *S* was freshly diapered, had just eaten to satiety, and had been awake for from 15 to 20 minutes. During the operant and extinction periods, at the end of every five-minute interval, *S* was placed in a prone position to rest for three minutes. While he was being turned over and during the rest period, *E* was not in his field of vision. During the experimental period, *S* was frequently required to work without rest for two successive intervals, but not for longer.

Typically, there were two to three sessions per day. Length of any one session ranged between 10 and 60 minutes, and was a function of the length of time *S* remained awake. The total number of days spent in experimenta-

tion with any one *S* ranged from eight to 16. Because of the very large amount of time per day required for travel, experimentation, and maintenance of rapport with the mother, only one *S* was run at a time.

Results and Discussion

Performance During the Operant Period

The two groups did not differ significantly in either mean operant response rate ($t = .29$, 7 *df*) or total number of responses emitted ($t = .06$, 3 *df*). For this reason, scores have been combined for graphic presentation (Fig. 1).

Figure 1 also includes a cumulative plot for a ninth, control *S*, who was run, without reinforcement, for an extended operant period of 19 intervals—or roughly three times the length of the operant period for experimental *S*s. These data provide some evidence against the possible explanation that the mere presence of *E*—and not the reinforcement—is conducive to increased rate of response.

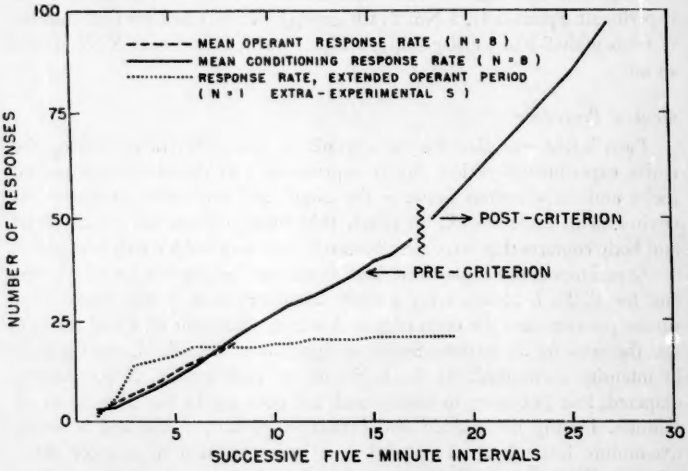

Figure 1—Cumulative curves showing rate of smiling response during operant and conditioning periods.

Performance During the Conditioning Period

For the conditioning period, the total number of responses for each group was 927, while the mean number of responses per subject was 231.7. Mean group response rates are presented in Table 1; the most noticeable

features of these data are the abruptness of the change in mean response rate and in variability following criterion.

Relative stability in response rate is also reflected in Table 2, which contains the individual post-criterion response rates emitted under the regular reinforcement schedule. For the IR *Ss*, percentages are based on the 10 intervals of regular reinforcement following criterion. For the RR *Ss*, percentages are based on the total number of intervals following criterion. From this table it can be seen that although most *Ss* habitually worked at one rate of response, there were interindividual differences in the habitual or "preferred" rate.[6]

TABLE 1

GROUP RESPONSE RATES DURING OPERANT AND CONDITIONING PERIODS

	Operant Response Rate		Conditioning Response Rate					
			Preceding Criterion		For First 10 Intervals Following Criterion		For 11th–22nd Intervals Following Criterion	
	Mean	σ	*Mean*	σ	*Mean*	σ	*Mean*	σ
Regular Group . .	2.11	.63	2.70	.47	5.15	.06	5.15	.06
Intermittent Group	2.88	.36	2.43	.44	5.10	.06	6.32*	1.11*
							8.12†	.64†
							13.00‡	1.51‡
Combined Groups	2.49	.64	2.56	.50	5.12	.06		

* 2:1 schedule.
† 3:1 schedule.
‡ 4:1 schedule.

Performance During the Extinction Period

Two statistical tests were applied to the extinction data in order to test the hypothesis that the IR group would be more resistant to extinction. First, McNemar's pseudo three-way analysis of variance (5), with blocks representing the experimental variable of reinforcement schedule, was applied to the response frequencies over all 13 extinction intervals. (Summing

[6] One last datum concerning performance during conditioning was provided by *S* No. 8, who gave a clear demonstration of the type of response that Skinner has termed "superstitious behavior" (6). The behavioral sequence was as follows. During reinforcement, *S* kept his left fist doubled in his mouth. When placed in the crib, he withdrew the fist from the mouth and kept it suspended in air for the short time it took him to smile. Then, simultaneously with the beginning of the reinforcement procedure, the fist was promptly reinserted into the mouth, the head turned 90 degrees to the left, and the body musculature stiffened. The onset of this stereotyped response coincided with criterion (56th trial). It disappeared for three days during a period of illness (90th through 118th trials), reappeared in full strength with recovery (119th trial), gradually diminished in intensity, and finally disappeared altogether by the 162nd trial.

TABLE 2

INDIVIDUAL RESPONSE RATES DURING CONDITIONING: PERCENTAGE OF INTERVALS FOLLOWING CRITERION DURING WHICH *S* EMITTED FOUR, FIVE, OR SIX RESPONSES PER INTERVAL

	Total Number of Intervals on Regular Reinforcement Following Criterion	*Percentage of These Intervals During Which Response Rate Was*		
		4	5	6
Intermittent Group				
S No. 1	10	10%	80%	10%
S No. 3	24	20	70	10
S No. 5	10	30	10	60
S No. 7	45	20	40	40
Regular Group				
S No. 2	10	4.2	58.3	37.5
S No. 4	31	13.3	62.2	24.4
S No. 6	10	80.6	19.4	0
S No. 8	22	0	22.7	77.3

across 13 intervals, the total number of responses for the IR group was 331; for the RR group, 130.) The resulting F of 17.14, with 1 and 6 df, is significant beyond the .01 level. Second, a mean difference was computed, using the four matched-pair differences in total number of responses during the last six extinction intervals. (In this case, total number of responses for the IR and RR groups were 113 and 3, respectively.) For 3 df the t of 6.77 is significant beyond the .005 level by a one-tailed test.

A noncumulative plot of mean response rates during extinction (Fig. 2), shows some interesting periods of sharp rise in response rate for both groups (interval No. 5 for the RR group; intervals No. 7 and No. 8 for the IR group). The point of occurrence of this "recovery" did not appear to be a function of the length of preceding rest period; e.g., in the IR group, two *S*s recovered after a rest of several hours (i.e., at the beginning of a new session), while the other two *S*s recovered after a rest of only three minutes (i.e., during a session).

Figure 2 also indicates that every member of the RR group extinguished not to his previous, operant rate of response, but to a zero rate. Coincident with the beginning of zero response rate was a conspicuous behavioral change: *S* would no longer fixate the discriminative stimulus (*E*'s face). Instead, *S*'s head turned to one side and remained there—an occurrence, it might be pointed out, that was in distinct contrast to *S*'s persistent fixation during conditioning. When this occurred, immediately preceding the last extinction interval, *E* propped *S*'s head with rolled blankets or other material, making it impossible for the infant to turn his head to the left

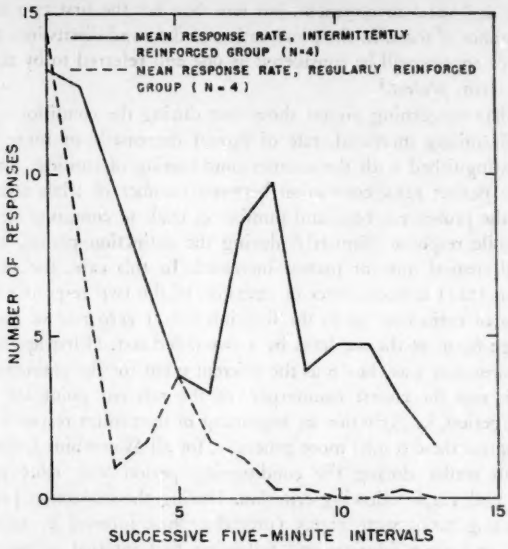

FIGURE 2—Noncumulative curves showing mean rates of smiling response during extinction period.

or right more than a few degrees. The "refusal" to fixate persisted even under these conditions; *S*'s eyes then turned toward the ceiling. When withdrawal of reinforcement is conceptualized as frustration-producing (1), this persistent nonfixation may be regarded as an avoidance response that is elicited by continued frustration of the original (approach) response, is incompatible with the original response, and is reinforced by repeated escape from the frustrating situation.

The Relation of Protest to Smiling

As mentioned above, one of the criteria for subject selection was the frequency and intensity of crying. This criterion necessitated continued notations of such behavior during the operant period and first few conditioning intervals. Even after this point, however, *E* continued to record the incidence of crying, although at the time there seemed to be no particular reason for continuing to collect these data.

In recording crying, one of three types of notation was made, according to the intensity of response. In decreasing order of intensity, the notations were: (a) crying; (b) intense fussing: the same type of muscular and vocal involvement as in crying, but tears appeared only at the corners of eyes and did not course down the cheeks or temples; (c) fussing: considerable

muscular and vocal involvement, but less than for the first two categories; no appearance of tears. In the presentation of data and discussions to follow, all three categories will be considered as one and referred to by the generic response term, *protest*.[7]

The data concerning protest show that during the conditioning period, as rate of smiling increased, rate of protest decreased—or more properly, protest extinguished with the counter-conditioning of smiling. Specifically, there is a perfect rank correlation between number of trials taken to extinguish the protest response and number of trials to conditioning criterion on the smile response. Similarly, during the extinction period, as rate of smiling decreased, rate of protest increased. In this case, the rank order correlation (tau) between rates of emission of the two responses from the beginning of extinction up to the first interval at zero rate of response, is —.69, significant at the .02 level by a two-tailed test. (First appearance of zero response rate was chosen as the referent point for the extinction period because it was the closest counterpart of the referent point for the conditioning period, i.e., criterion or beginning of maximum response rate.)

To express these results more generally, for all *Ss* combined, the ratio of protests to smiles during the conditioning period was 1:6.5 preceding criterion, and 1:276 following criterion. During the extinction period, the corresponding ratios were 1:40.5 (preceding first interval at zero rate of response) and 1:2.7 (during and following first interval at zero rate of response). Rates of protest response during operant, conditioning, and extinction periods are shown in Figure 3.

One might conceivably object to any interpretation of these correlations on the grounds that they merely reflect the fact that two such responses are mutually exclusive behaviors: an infant cannot protest and smile at the same time. Although this overlooks the important point that all *Ss* ***changed*** in the frequency of emission of both responses, there is a direct answer to such an objection in that the precriterion data show, for every *S*, both part and whole intervals during which neither protest nor smiling occurred at all.

There is a good deal of similarity between these findings and results obtained in a study by Estes (3), in which rats were reinforced for one instrumental response while a previously reinforced, competing response was being extinguished. Estes states that, "These results seem to lend some support to the view that the 'learning' of any one response involves the concurrent extinction of others and that the amount of initial acceleration

[7] It should be noted that interjudge reliability in discriminating protest from nonprotest was not determined. However, as indirect support for the objectivity or reproducibility of these data, the following points should be considered. First, this was a simple discrimination to make; even responses of lowest intensity represented a marked contrast to the infants' typical placid behavior. Second, the total numbers of protests for the two groups were approximately equal. Third, the orderliness of these data was not observed until the end of the experiment, when the results were actually tabulated.

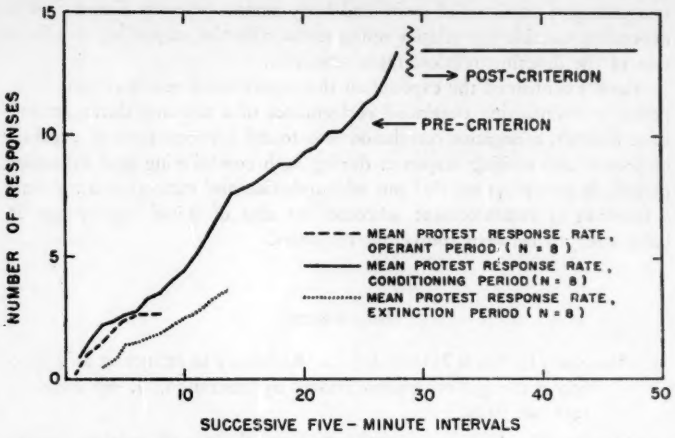

FIGURE 3—Cumulative curves showing mean rates of protest response during operant, conditioning, and extinction periods.

in a learning curve is determined to an important extent by the relative initial strengths of all behaviors which may occur in the experimental situation" (3, p. 204).

In the present case, the most important determiner of rate of acquisition of the smiling response appears to be the initial strength of the functionally equivalent, competing protest response. The main difference between procedures is that in Estes' study, the conflicting response was both reinforced and extinguished as an integral part of the experiment, while in the present study only the extinction of the competing response took place during the experiment. Its establishment had already taken place prior to the beginning of the experiment—most probably via reinforcement by maternal care.

Perhaps the most appropriate way of conceptualizing these results is in terms of a habit family hierarchy of two responses for which the initially differing habit strengths were first reversed by selectively reinforcing only the weaker response, and then reversed again by extinguishing that response, allowing for recovery of the first. This leads to the general proposition that rate of acquisition and extinction is a function of the initial discrepancy in habit strength between competing responses.

SUMMARY

This investigation was concerned with the instrumental conditioning of a social response (smiling) in infants. *Ss* were two groups of four infants each; one group was maintained on a conditioning schedule of intermittent reinforcement and the other on a schedule of regular reinforcement. The

reinforcement consisted of social and body contact between *E* and *S*. The dependent variable was relative resistance to extinction of smiling as a function of the differing reinforcement schedules.

Results confirmed the expectation that intermittent reinforcement is superior in maintaining continued performance of a response during extinction. Further, a negative correlation was found between rates of emission of protest and smiling responses during both conditioning and extinction periods. It was proposed that rate of acquisition and extinction is not only a function of reinforcement schedule but also of initial discrepancy in habit strength between competing responses.

References

1. Adelman, H. M., & Maatsch, J. L. Resistance to extinction as a function of the type of response elicited by frustration. *J. exp. Psychol.*, 1955, 50, 61-65.
2. Dennis, W. An experimental test of two theories of social smiling in infants. *J. soc. Psychol.*, 1935, 6, 214-223.
3. Estes, W. K. Effects of competing reactions on the conditioning curve for bar pressing. *J. exp. Psychol.*, 1950, 40, 200-205.
4. Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: A review and critique. *Psychol. Bull.*, 1950, 47, 193-234.
5. McNemar, Q. *Psychological statistics.* New York: John Wiley, 1955.
6. Skinner, B. F. "Superstition" in the pigeon. *J. exp. Psychol.*, 1948, 38, 168-172.
7. Thompson, Jane. Development of facial expression of emotion in blind and seeing children. *Arch. Psychol.*, 1941, 37, No. 264, 1-47.
8. Washburn, Ruth. A study of the smiling and laughing of infants in the first year of life. *Genet. Psychol. Monogr.*, 1929, 6, 397-537.

COMPARATIVE SOCIAL ADJUSTMENTS OF ELEMENTARY SCHOOL PUPILS WITH AND WITHOUT PRESCHOOL TRAINING

MERL E. BONNEY
North Texas State College

ERTIE LOU NICHOLSON
Department of Research
Public Schools of Dallas, Texas

Educators, psychologists, and sociologists have generally agreed upon the importance of early childhood in the formation of personality and character. No doubt this emphasis is well placed. However, there is still much to be found out about the conditions or kinds of experiences which make for significant differences in subsequent development. It is the major purpose of this paper to present a report of some efforts to evaluate the extent to which nursery school and kindergarten experiences can be shown to make a difference in classroom social adjustments in subsequent elementary school grades. The primary question to be answered is: Can those children, as a group, who have had these preschool socializing experiences be shown to have measurable advantages in subsequent years in classroom social adjustments as compared with children in the same classes who have had no nursery school or kindergarten training?

This report consists of a review of three studies bearing on this question, together with a discussion of the findings.

FIRST STUDY

Methods and Subjects

The first study to be reported was conducted by Angell (2) as a part of her graduate work at North Texas State College. Her subjects were 78 children in two kindergartens, two first grades, one second grade, and one third grade in two schools in Denton, Texas. These children were

divided into 39 pairs with one pupil of each pair having attended a nursery school while the other member of each pair had not.

In order to help equate these pairings the two pupils put together within each class were similar in sex, fathers' occupational level, and number of siblings in the family. Classification of fathers' occupations was based on the Revised Scale for Rating Occupations as given in Warner, Meeker, and Eells (14).

All of the children in the above listed grades responded to a one-criterion sociometric test involving several choices per pupil for preferred playmates. Total scores were calculated for each pupil on the basis of this measurement.

The teachers in four of the classes, including kindergarten through third grade, rated all the pupils included in their respective classes on the Winnetka Scale for Rating School Behavior and Attitudes (now out-of-print). This scale has twelve subdivisions, but the ratings obtained on all of these are summarized into five major trait-categories as follows: Co-operation, Social Consciousness, Emotional Adjustment, Leadership, and Responsibility. The teachers did not know the purpose of this study when they made their ratings.

Findings

The difference between the mean sociometric scores of the nursery and non-nursery school pupils was found by *t* test to be significant at the .02 level of confidence.

In the four classes in which teacher ratings were obtained there were 23 of the 39 pairings described above. When the mean score of the Winnetka Scale in each of the five major categories for the 23 pupils who had attended nursery school was compared with that for the 23 non-nursery school pupils, only one of the differences was found to be statistically reliable, this at the 5 per cent level for the trait designated Social Consciousness.

Summary

These findings clearly show that the pupils who had attended nursery school had a reliable advantage over the non-nursery school pupils in receiving positive choices from their classmates, but they showed a reliable advantage in only one out of five traits from the standpoint of teacher ratings on the Winnetka Scale.

Second Study

Methods and Subjects

The second study was a master's thesis completed at North Texas State College by Nicholson (10). Her subjects were 402 children in grades one through six in a public elementary school in a predominantly lower-middle class neighborhood in Dallas, Texas.

The first step in this study was to send a questionnaire to all the parents of all the pupils in this school asking them to supply information on whether or not their children had attended a nursery school or a kindergarten, or both. Of nearly 1000 questionnaires sent out, approximately 900 were returned. From these returns it was possible to select 14 classes (out of 32 classes available in the school) in which approximately 50 per cent of the pupils had had some type of preschool training. There were two classes on each of four grade levels and three classes on each of the other two grade levels.

Of the 402 pupils included in this study a little over 200 had had no formal preschool training. Of the nearly 200 who had had some type of preschool experience, approximately one-fourth had attended a nursery school, about one-half had attended a kindergarten, and the remaining one-fourth had attended both a nursery school and a kindergarten.

Classroom social adjustment was measured by sociometric testing and by teacher nominations.

In each of the 14 classes the pupils made four choices for others in their respective rooms with whom they would most like to eat lunch, and also four choices for the pupils they would select as candidates for the "Best Citizen of the Year" award—an annual custom in this school. These two criteria were used because they were both appropriate to all six grade levels in this school. Each child's total choice-status was based on his combined score from both of these criteria.

Teacher nominations were obtained by having each teacher list the upper one-fourth of pupils in her class who were regarded by her as making the most effective and satisfactory social adjustments within her classroom group; also, each teacher listed those whom she considered to be in the lowest fourth in reference to classroom social adjustments. All these listings were made before the teachers knew anything about the purposes of the investigation.

Findings

One of the major hypotheses which was tested in this study was whether the factor of preschool attendance or nonattendance was related to being in the upper or the lower halves of the sociometric distributions of the respective grade groupings. In order to have sufficient cases for statistical treatment the several classes on each grade level were combined, but all grade-level groupings were treated separately. The number of cases on the six grade levels ranged from 49 to 85 with a median of 67.

Using four-fold contingency tables each category was then compared with every other category. Because of the small number of cases in some cells an allowance for discontinuity was made by applying Yates' correction in computing the chi squares. When these data were assembled, it was found that none of the chi squares was statistically significant. These

findings obviously do not support the hypothesis that there was a significant relationship between choice-status, as measured in this research, and preschool attendance or nonattendance.

A second hypothesis tested in this investigation was that the teachers' listings of pupils in the upper and lower fourths from the standpoint of classroom social adjustments, as described above, would show a reliable relationship to the fact of preschool attendance or nonattendance. This hypothesis was also tested by chi square, using Yates' correction. None of these relationships proved to be even close to statistical significance.

A third hypothesis bearing on the carry-over value of preschool training was tested by forming matched groups. On each grade level those pupils who had attended some type of preschool were matched with nonattending pupils in regard to the following factors: sex, age, fathers' occupational classification, and IQs obtained from a group-administered intelligence test. On the basis of this matching the following number of pairs was obtained on each grade level from the first through the sixth: 23, 33, 34, 24, 20, 28. When the *t* test was applied to each of these experimental and control groupings, as they were related to sociometric choice-status on each grade level, none of the differences was found to be statistically significant. Only one of them even approached the .05 level of confidence.

Summary

It is clear that these findings provide no evidence that those pupils who had attended some type of preschool had, as a group, any advantage in personal-social behavior over those who had had no preschool training, either from the standpoint of acceptability by classroom peers or on the basis of teacher evaluations.

Third Study

Methods and Subjects

The third study was concerned with four sixth grade classes in the Denton, Texas, public schools. The pupils in each of these four sixth grades responded to a two-criteria sociometric test involving choices for preferred playmates and for desired associates in classroom group work. Also, each of the four teachers listed the names of those pupils whom she considered to be in the upper fourth of her class in "over-all good adjustment to the classroom social situation," and those she considered to be in the lowest fourth of her group on the same criterion.

Findings

In three of the four sixth grades approximately 35 per cent of the pupils had attended either nursery school or kindergarten. In these classes the average number of choices received on the two sociometric criteria combined was very nearly the same for the preschool and the non-preschool

groups. The actual difference was less than one point in each of the three classes.

In the remaining sixth grade 80 per cent of the pupils had attended either nursery school or kindergarten, or both. In this class the preschool pupils received an average sociometric score which was nearly twice as large as that received by the non-preschool pupils. This is obviously a marked advantage but probably should not be used as a basis for generalization because of the very unbalanced proportions of the two groups being compared in this class.

If the analysis of teacher placements into upper and lower fourths, as described above, is confined to the three classes in which the number attending some type of preschool amounted to approximately 35 per cent of each class, the conclusion is reached that these placements were unrelated to the factor of preschool attendance or nonattendance. Both the upper and lower fourths in each class included preschool and non-preschool pupils in numbers which were very close to what would be expected from their proportional representation in their respective groups.

Considering all four of the sixth grades, an effort was made to see if those pupils who had attended *both* nursery school and kindergarten would show up better on the sociometric testing and the teacher placements than those who had not attended either type of preschool. In all four classes combined there were 20 pupils who met this condition. An inspection of their average choice-status in their respective classes and their listings on the teacher placements failed to show any advantage for these pupils except in the room mentioned above in which only 20 per cent of the children had had no preschool training.

An effort was made to isolate the influence of the factor of preschool vs. non-preschool attendance by comparing these two groups in each class on the basis of IQs, fathers' occupational classification, and family size. The IQs were obtained by administering the California Short-Form Test of Mental Maturity. The fathers' occupational classifications were based on the same scale used in the first study. Family size was based on the number of siblings within six years of the age of each pupil being studied.

In none of these three variables was the difference between the preschool and the non-preschool groupings large enough to warrant statistical treatment. All the average IQ differences were less than seven points, and in three of the four classes the differences were only two or three points. All the differences between average scores for fathers' occupational classifications were less than one point with the exception of the class in which only 20 per cent of the pupils had not gone to preschool. In this class those who had attended preschool possessed an obvious advantage in fathers' occupational level.

In regard to family size the means and medians for all the groups were very similar, the largest single difference being less than one-half of a point.

Summary

From the foregoing data it is clear that the over-all picture is one of no significant differences between the preschool and non-preschool pupils in social adjustments in the sixth grades studied.

Review of Related Research

Other studies which have been designed to evaluate preschool education, particularly the nursery school, have shown some definite advantages for this type of education, although these advantages have not always been very outstanding or consistent on various types of measurements.

Kawin and Hoefer (7) have reported an investigation involving comparisons between 22 children who attended a nursery school and 22 who did not attend. No differences were found in physical characteristics or in degree of mental growth, but some differences were found in favor of the nursery school pupils in eliminating certain undesirable habits and in developing more independence of adults.

In a more extensive study Hattwick (5) compared ratings obtained by three nursery school teachers on 106 pupils who had been in nursery school attendance for nine months with an equal number who had been in attendance in the same schools only six weeks and who were matched with the first group on age, sex, nationality and race, and economic level. On most items involving social behavior and routine adjustments (but not emotional traits) the more experienced nursery school pupils were reliably superior to those who had been in attendance only six weeks.

Other similar studies by Joel (6), Parten (8), and Walsh (13) have also shown some beneficial results in various kinds of personal and social behavior from nursery school attendance.

On the other hand, Andrus and Horowitz (1) found little evidence that length of time spent in a nursery school increased feelings of security when these feelings were measured by three adult raters. The authors, however, were quite critical of the rating scale which they used. In a similar vein, Cushing (3) found "no striking differences" in personality trait ratings as given by nine kindergarten teachers to 33 pupils who had previously attended a nursery school as compared with a corresponding group who had not attended. Likewise, Phillips (9) found no advantage among first graders on the California Test of Personality for pupils who had had kindergarten training over those not having had such training.

Discussion

One question which would naturally be raised in a report of this kind is in regard to the *quality-level* of the preschool training which any particular group of pupils receives. The only basis available for dealing with this point in reference to the three studies which constitute the main

body of this report is the proportion of subjects that attended a nursery school or a kindergarten sponsored by one of the two state colleges in Denton, the assumption being that these preschools are probably representative of the best kind of training available.

In the first study conducted by Angell nearly all the pupils had had their training in a college-sponsored preschool; in the second study probably none of the pupils had attended a college-sponsored preschool; while in the third study about 40 per cent of the sixth graders who had had preschool training had attended one of the college-sponsored nursery schools or kindergartens.

The one finding which showed the most clear-cut advantage for the preschool pupils was from Angell's study in which a very high percentage of the pupils had attended a college-sponsored preschool, and was based on sociometric data. However, teacher ratings in four of the six classes in this study showed only a small advantage for those who had attended these college-sponsored preschools.

Data from the four sixth grades in the third study do not lend support to the idea that attendance at a preschool of high quality may be reflected in better personal-social adjustment over a period of five or six years; but these data have only limited relevance to this point since not more than 40 per cent of the preschoolers in these classes had attended a college-sponsored nursery school or kindergarten.

One experimental study by Thompson (11) is especially pertinent to the matter of quality-level since his evidence showed that nursery school pupils whose teachers manifested toward them a large amount of warm friendly and helpful relationships made much more progress in various aspects of personal-social behavior than did a similar group of nursery school pupils whose teachers were considerate but who were somewhat detached and gave help only when it was specifically requested. These findings suggest that one of the essential characteristics of a preschool of high-quality level is a large amount of personal-social interaction not only between the pupils but also between the pupils and their teachers. Unfortunately, information is lacking on the extent to which this kind of relationship was present in the great majority of preschools that have been studied.

Another factor or condition pertinent to the present discussion is the extent to which the benefits of nursery school traning can be shown to carry over into subsequent years, a determination of which was the main purpose of the three studies reported here. It would seem that this type of information is especially needed to evaluate preschool education since most of the studies previously reported have been confined to evaluating preschool experiences only during the time the preschool has been in session, with no subsequent follow-up. This is true of all of the studies reviewed or referred to above with the exception of the ones by Cushing and by Phillips and neither of these showed any advantage for preschool training. Without some efforts to trace the effects of preschool education

into subsequent years there is not an adequate basis for evaluating the contributions of special training as compared to simply general development.

One extensive investigation which has a bearing on this topic has been reported by Van Alstyne and Hattwick (12). They conducted a follow-up study into the elementary school grades of 165 children who had had nursery school training. One of the outstanding findings was a very strong tendency for those who were classified as well adjusted or as poorly adjusted in the nursery school to be likewise classified in their subsequent elementary school classes. The authors state: "The whole trend of the results . . . shows how closely behavior in the nursery school parallels that of the upper grades" (p. 67).

This finding, together with a corresponding finding in a similar study by Fisher (4), emphasizes the persistence of a certain level of personal-social adequacy irrespective of group socialization experiences.

These results, together with the evidence presented above in the three studies from the North Texas area, lead to the conclusion that if early socialization experiences are going to possess significant carry-over values into subsequent years they probably will need to be of a particularly high quality-level in reference to interpersonal rapport between the pupils and their adult supervisors and also in regard to adequate provision for meeting the varying needs of individuals.

It seems likely that many educators and child psychologists have had too much naive faith in the adjustment or curative values of group socialization, without sufficient attention being paid to whether or not the activities engaged in are actually helping those who need some assistance as opposed to simply perpetuating an established social heirarchy.

If this is a valid statement, then it follows that nursery school and kindergarten teachers should become more critical of their programs, particularly from the standpoint of planning more exercises, both in the school and with parents, which are designed to help those who show obvious signs of personal and social maladjustments.

References

1. Andrus, R., & Horowitz, E. L. The effect of nursery school training: insecurity feelings. *Child Develpm.*, 1938, 9, 169-174.
2. Angell, Dorothy B. Differences in social behavior between elementary school children who have attended nursery school and those who have not attended nursery school. Unpublished master's thesis, North Texas State College, 1954.
3. Cushing, H. M. A tentative report on the influence of nursery school training upon kindergarten adjustments as reported by kindergarten teachers. *Child Develpm.*, 1934, 5, 304-314.

4. Fisher, M. L. Measured differences between problem and non-problem children in a public school system. *J. educ. Sociol.*, 1934, 7, 353-364.

5. Hattwick, B. W. The influence of nursery school attendance upon the behavior and personality of the pre-school child. *J. exp. Educ.*, 1937, 5, 180-190.

6. Joel, W. The influence of nursery school education upon behavior maturity. *J. exp. Educ.*, 1939, 8, 164-165.

7. Kawin, E., & Hoefer, C. *A comparative study of a nursery school versus a non-nursery school group.* Chicago: Univer. of Chicago Press, 1931.

8. Parten, M. B. Social participation among pre-school children. *J. abnorm. soc. Psychol.*, 1932-1933, 27, 243-269.

9. Phillips, G. J. Kindergarten training as a factor in the social adjustment of a selected group of young children. Unpublished master's thesis, Eastern New Mexico Univer., 1953.

10. Nicholson, Ertie L. The relative social development of children with pre-school background as opposed to those who lack such experiences. Unpublished master's thesis, North Texas State College, 1957.

11. Thompson, G. G. The social emotional development of pre-school children under two separate types of educational program. *Psychol. Monogr.*, 1944, 56, No. 5.

12. Van Alstyne, D., & Hattwick, L. A. A follow-up study of the behavior of nursery school children. *Child Develpm.*, 1939, 10, 43-70.

13. Walsh, M. E. The relation of nursery school training to the development of certain personality traits. *Child Develpm.*, 1931, 2, 72-73.

14. Warner, W. L., Meeker, M., & Eells, K. *Social class in America.* Chicago: Science Research Associates, 1949.

A SUGGESTED REVISION OF NORMS FOR THE *ROGERS TEST OF PERSONALITY ADJUSTMENT* [1]

LEE G. BURCHINAL, BRUCE GARDNER, *and* GLENN R. HAWKES
Child Development Research Laboratories
Iowa State College

Although it is an old test, the Rogers Test of Personality Adjustment is still used in child guidance and school guidance programs. Yet, there have been few subsequent publications relating to this test since Rogers' original report (5). Recently this test was used in a study based on a sample of fifth grade children. The results of this study provided an opportunity to determine whether the norms given by Rogers in 1931 are still adequate for use today.

METHODOLOGY

The data for the present study are based upon completed Rogers tests from 256 fifth grade children selected from rural areas and small towns (less than 10,000 population by the 1950 Census) in Iowa, Ohio, Kansas, and Wisconsin. Eight sampling points were selected for each state sample by a probability method. At each sample point, eight children were randomly selected from the fifth grade. All the children came from "whole" families, ones in which both parents were living together, and all the children had at least one brother or sister. Field work was completed during the 1954-1955 school year. Detailed descriptions of the sample design and the children's family characteristics have been reported elsewhere (1, 2, 3).

[1] Published as Journal Paper No. J-2963 of the Iowa Agricultural Experiment Station, Ames, Iowa, Project No. 1171, Home Economics Research. Acknowledgment is due Leone Kell of the Kansas Agricultural Experiment Station, Ruth Hoeflin of the Ohio Agricultural Experiment Station, and Helen Dawe of the Wisconsin Agricultural Experiment Station for their collaboration in this regional research project.

Since descriptions of the Rogers Test of Personality Adjustment are available (4; 5, pp. 354-356), only a brief discussion is given here. Four subscores, personal inferiority, social maladjustment, family relations, and day dreaming, are derived from the test. When 43 children were retested by Rogers after an interval of one month, subscore reliabilities ranged from .65 to .72. The total score reliability was .72. In the present study, 51 children were retested after a one-week interval, and subscore reliabilities varied from .67 to .77. The total score reliability was .71. The results of the several efforts made by Rogers to test the validity of the adjustment tests were accepted for the purposes of the present study.

Rogers developed a set of norms for use with his test based on a total of 167 children. In developing his norms, Rogers used clinical data as well as means and standard deviations. Taking these factors into account, Rogers divided the scores into three groups, high, average, and low levels of maladjustment, because he felt finer gradations were not warranted on the basis of his data. The extreme categories approximated quartile groupings of scores. For Rogers' norms, see Table XX (5, p. 86).

TABLE 1

PERCENTAGE OF LOW, AVERAGE, AND HIGH SCORERS ACCORDING TO ROGERS' NORMS

	Score Classification		
Adjustment Scores	*Low*	*Average*	*High*
Personal Inferiority	52	28	20
Social Maladjustment	12	36	52
Family Relationships	27	41	32
Day Dreaming	38	29	33
Total Score	23	51	25

The Findings

When the scores of 256 children tested in this study were classified according to Rogers' norms for the four subscores and the total scores, the results presented in Table 1 were obtained.

It should be noted that on the personality inferiority scale 52 per cent of the children were low scorers while on the social maladjustment exactly the same percentage were high scorers. It would seem that any set of norms that places over half of the cases into one or the other extreme scoring category is unrealistic. The percentage distributions for the family relationships scores approached what might have been expected, while the distribution of the total scores followed closely a quartile distribution. The distribution for the day dreaming scores was bimodal with the smallest frequency in the "average" category.

In order to test whether the frequencies for the percentages listed in Table 1 differ significantly from reasonable estimates of the distributions of the scores, an expected distribution for the scores was defined by using quartile groupings of scores. Except for the absence of clinical data, this method of defining low, average, and high scoring categories approximated a replication of the procedure followed by Rogers. Each of the extreme categories should have included approximately 64 cases while the average category should have included 128 cases. Chi square comparisons were used to test the departure of the observed from the expected frequencies. For simplicity, all the computation tables have been omitted.

The distribution for the total scores was the only one which did not depart significantly from the expected distribution ($.50 < p < .70$). The chi squares for the four subtests were significant ($p < .01$).

Since the observed frequencies for the classification of scores departed significantly from what might have been expected on all four of the subscores, a new set of norms based on the scores of the 256 children tested in this study was developed. Differences between boys' and girls' mean scores were tested first to determine if separate norms ought to be used for some of the scores. Mean scores and standard deviations for boys and girls and for the total sample for the five scores are given in Table 2.

TABLE 2

MEAN SCORES AND STANDARD DEVIATIONS FOR BOYS, GIRLS, AND TOTAL SAMPLE ON ROGERS TEST

	Boys $N=129$		Girls $N=127$		Total Sample $N=256$	
	M	*SD*	*M*	*SD*	*M*	*SD*
Personal Inferiority ...	10.81	3.95	12.43	4.99	11.61	4.57
Social Maladjustment ..	14.74	5.74	14.85	4.07	14.80	4.43
Family Relations	9.18	3.88	8.72	3.78	8.95	3.84
Day Dreaming	2.88	2.28	3.31	2.77	3.09	2.55
Total Score	37.60	8.00	39.17	8.39	38.38	8.23

Since the boys' and girls' means were significantly different on the personal inferiority scale, it seemed that separate norms should be considered for these scores; otherwise, similar norms were developed for boys and girls for the other four scores.

Suggested norms for the Rogers scores are listed in Table 3. The cutting points for the low, average, and high scoring groups were determined by taking the scores which came the closest to the dividing lines for the top and bottom quartile.

The modified norms differ considerably from Rogers' norms for the personal inferiority and social maladjustment dimensions of the adjustment test. Also, separate norms for boys and girls are suggested on the personality inferiority scores. Only small changes occurred in relation to the family relationships scores. Several changes were observed for the day dreaming scores. Rogers suggested separate norms for boys and girls on this dimension of his test. The present data suggested separate norms were not necessary. The new norms did not conform to either of the sets given by Rogers, but approximated the averages of the separate norms given by him. Finally, no changes were necessary for the total score norms.

TABLE 3

SUGGESTED NORMS FOR THE ROGERS TEST

	Score Classification		
Adjustment Scores	*Low*	*Average*	*High*
Personal Inferiority			
Boys	− 8	8–13	14+
Girls	− 9	9–15	16+
Social Maladjustment	−12	12–17	18+
Family Relations	− 7	7–11	12+
Day Dreaming	− 2	2– 4	5+
Total Score	−33	33–43	44+

SUMMARY

Data derived from the administration of the Rogers Test of Personality Adjustment to a sample of 256 fifth grade children provided a basis for evaluating the norms given by Rogers in 1931. Rogers' norms appeared to be inadequate for classifying children into low, average, and high scoring categories for all the subscores of the test. The norms for the total scores are apparently still adequate.

When it was apparent that Rogers' norms were not realistic for present use of the test, boys' and girls' mean scores were compared to determine if separate or similar norms should be used for them on the various Rogers scores. The means of the two groups were not significantly different except for the personal inferiority scores. Therefore, boys and girls were separately classified for the development of norms for the personal inferiority scores. The total sample of all children was used for the other scores. Scores corresponding to the cutting points for the top and bottom quartiles were used to define the low, average, and high quartiles. On this basis, a new set of norms was proposed.

References

1. Burchinal, L. G. The relations of parental acceptance to adjustment of children. Unpublished doctor's dissertation, Ohio State Univer., 1956.

2. Burchinal, L. G., Hawkes, G. R., & Gardner, B. The relationship of parental acceptance to the adjustment of children. *Child Develpm.*, 1957, 28, 65-77.

3. Hawkes, Glenn R., Burchinal, L. G., Gardner, B, & Porter, B. M. Parents' acceptance of their children. *J. Home Econ.*, 1956, 48, 195-200.

4. Remmers, H. H., & Gage, N. L. *Educational measurement and evaluation.* New York: Harper, 1943.

5. Rogers, C. *Measuring personality adjustment in children nine to thirteen years of age.* New York: Teachers College, Columbia Univer., 1931.

RESEARCH ON THE EFFECTS OF NON-CONTINUOUS MOTHERING[1]

DAMARIS PEASE *and* D. BRUCE GARDNER
Child Development Research Laboratories
Iowa State College

That there is some significant relationship between the development of the infant and the nature of his early mothering experiences would not appear to be open to serious question at this point. The negative effects of inadequate mothering in the case of the severely deprived infant have been repeatedly observed by a number of research workers (2, 5, 9, 10, 12, 15, 16, 17, 19, 20).

Bowlby (5), in a summary of such research, considers it highly significant that different research workers, operating independently and to some degree in ignorance of each others' efforts, arrived at essentially similar conclusions on the adverse effects of mother-child separation.

A clinical description of the deprived infant, which for all its brevity is remarkably comprehensive, is presented by Bakwin:

> Infants under 6 months of age who have been in an institution for some time present a well-defined picture. The outstanding features are listlessness, emaciation and pallor, relative immobility, quietness, unresponsiveness to stimuli like a smile or a coo, indifferent appetite, failure to gain weight properly despite the ingestion of diets which, in the home, are entirely adequate, frequent stools, poor sleep, an appearance of unhappiness, proneness to febrile episodes, absence of sucking habits (3, p. 512).

Other writers (16, 20) have stressed the finding that institutional children in a physically sterile setting are more susceptible to colds and epidemics of infectious diseases than are children who have the psychological support of a family atmosphere.

Bowlby lists the typical characteristics of children deprived in infancy, when observed later in childhood, as follows: ". . . superficial relationships;

[1] Journal Paper No. J-3281 of the Iowa Agricultural and Home Economics Experiment Station, Ames, Iowa, Project No. 1299. Acknowledgment is due the Elizabeth McCormick Memorial Fund for partial financial support.

no real feeling—no capacity to care for people or to make true friends; an inaccessibility, exasperating to those trying to help; no emotional response to situations where it is normal—a curious lack of concern; deceit and evasion, often pointless; stealing; lack of concentration at school" (5, p. 31).

One is impressed, in reading such descriptions, that this wide range of symptoms represents not merely an isolated effect on some specific segment of personality, but a vast complex of responses of the total organism to an inadequate environment. It is conceivable that this complex of organismic responses, as the deprived individual moves forward over the time dimension, has its own pattern of lawful and orderly development, which we are only beginning to understand. Bowlby and Robertson, at least, are willing to assert that, "The general proposition that separation in the early years of life can have permanent adverse effects now belongs to history" (7, p. 123). They go on to say that we still need to know how these effects come about.

Goldfarb (15) presents a rather striking picture of the differences in behavior between children who were institutionalized as infants and children who were reared in foster homes. His comparisons were made between two small groups of children between the ages of 10 and 14, with respect to intelligence, ability to conceptualize, school achievement, social maturity, speech, response to frustration, and capacity for relationships. He concluded that the institutional group was significantly inferior to the control group in all of these characteristics. Comparison of Rorschach responses for the two groups led Goldfarb (12) to characterize the perceptions of the institutional children as looser, less well organized, confabulatory, and more arbitrary. He found lack of control over emotional responses for the institutional children, with evidence of diminished drive to social conformity.

To the extent that we can accept and generalize the findings of such research we might conclude that extreme and severe deprivation in early childhood initiates a progressive and irreversible developmental pattern. Further it would seem that the total organism is involved in a complex set of relationships in the progressive developmental pattern, and that the deprived child is especially characterized by reduced ability to utilize his personal and environmental resources, and lessened capacity for healthy affectional relationships with others.

The research on which such conclusions might be based has dealt largely with institutionalized infants and children, and has tended to make deprivation operationally synonymous with institutionalization. It becomes easy to think of the mother-child relationship as if it were a dichotomous variable, with deprivation either occurring or not occurring in any given situation. There is obviously no basis for making such an assumption; indeed it is probably safe to assume just the opposite, that there is infinite and continuous variation from some condition representing a most desirable mother-child relationship to a condition of complete deprivation. Although even these extremes are not precisely defined in the literature, it is likely

that considerable agreement could be found among professional people as to what constitutes the essential features of these conditions. But between these two relatively obvious extremes, it is less clear as to what constitutes intermediate amounts, kinds, and qualities of mothering, and what the effects of these relative degrees or conditions of mothering might be. Furthermore, it is a questionable assumption that only one continuum is involved. The complexity of factors suggests that the mother-child relationship has many dimensions, including that of time. That is, the nature of the conditions constituting adequate mother-child relationships may vary considerably with the age of the child. Thus, mothering itself, if defined operationally, must take into account the varying demands of a relationship which changes with time. Similarly, an operational definition of deprivation in all its infinite variations would necessarily take into consideration the notion that a given set of conditions might constitute varying degrees of deprivation at different times.

Goldfarb (15), in attempting to outline the differences between adequate and inadequate mother-child relationships, points out that in the ordinary family (a) the child has *specific* adults (parents) who are warm and loving; (b) the mother influence is *continuous,* in terms of daily routine and total life-span; (c) the mother is the source of safety and need-gratification to the child; (d) the child is given *constant stimulation* from his world of persons and objects; and (e) there is a reciprocal relationship between parent and child. The absence of these elements is considered maternal deprivation.

Even in the event of a clear-cut separation of the infant from its mother, there are obviously many conditions which may modify the effects of the experience on the child. Ainsworth and Bowlby comment on this:

> Both theoretical considerations and research to date suggest that separation experiences occuring under different conditions lead to effects which may differ qualitatively and/or quantitatively from one another. Certain of the conditions which seem likely to affect the outcome of separation . . . (include) . . . the state of psycho-social development and the strength of the dependency relationship which have been achieved by the time of separation and the nature of the relationship with substitute figures during the separation. Other conditions are: the degree of stability and security of the pre-separation dependency relationship with the mother; the length of the separation; the degree of continuity of the mother-child relationship during separation (for example through visiting); the nature of other relationships before and after separation, particularly the father-child relationship; the degree of previous familiarity with the separation environment and substitute figures, if any; the degree of frustration of other needs such as perceptual and activity needs during separation; and the reasons for separation and the child's understanding of them (1, p. 109).

The authors go on to suggest that we need to learn more about the role of such variables in modifying the effects of separation experience on the child. It is evident that there are a good many things we do not yet know, not only about the effects of separation and extremes of deprivation,

but about the basic elements of mother-child relationships which determine success, if the criterion for successful mothering is the optimum development of the child. The research of Dennis (8), which led to conclusions not entirely consistent with other research cited here, raises questions about the ostensibly deprived child whose development does not seem to be adversely affected by his experience. Stone points up our lack of understanding on such matters and urges cooperative research efforts:

> What, precisely, are the significant ingredients in the mothering relationship? Is it a specifically humanizing process, or to what extent is it essential to mammalian development in general? . . . What mothering practices are particularly harmful at which stages of development, and at which later stages of development will the damage or the benefit be evident? Here is opportunity for physician, experimental psychologist, clinical psychologist, case worker, and sociologist to join hands in order to arrive at some of the very fundamentals of the influences which shape and misshape human personality (23, p. 18).

Of special interest to the writers in their current research pertaining to non-continuous mothering has been the emphasis placed by Bowlby on the continuity of the mother-child relationship. At one point, in commenting on mental health, Bowlby defines mental ill-health as ". . . a derangement of the capacity to make co-operative relationships with other human beings (6, p. 117). Elsewhere Bowlby states: ". . . What is believed to be essential for mental health is that the infant and young child should experience a warm, intimate, and continuous relationship with his mother (or permanent mother-substitute) in which both find satisfaction and enjoyment" (5, p. 11).

Thus, it would appear that, to Bowlby at least, any mother-child situation in which this warm, intimate, *and continuous* relationship does not exist can be considered a state of maternal deprivation.

The research to be described here, currently in progress, represents one attempt to explore the continuity variable, and to determine what, if any, relationship might exist between the growth of the child and the continuity of his mothering experience in early infancy.

At certain colleges and universities the practice, in departments of home economics, has been to bring an infant into a home management house for a period of time to provide the students with an educational experience in the care of infants. Critics of this practice have argued that it constitutes a kind of maternal deprivation for the infant, since a number of women students, usually in rotation, have primary responsibility for the actual care of the infant. Thus, he is deprived of an experience with a single, continuous mother figure. A review of the literature fails to reveal any definitive statements of the effects on the child of being subjected to such a discontinuity of mothering.

Child Development and Home Economics personnel at Iowa State College have been interested in this question for a number of years. In

1955, in cooperation with the Iowa Children's Home Society, they sought and obtained financial support from the Elizabeth McCormick Memorial Fund to conduct research on the effects of home management house experience on the personality development of infants.

At the six home management house units currently operating at Iowa State College, the practice is to bring the infant into the house for a period of one college quarter, beginning usually when the infant is less than one month in age. Each house is equipped to accommodate a group of eight students and the house adviser. The students live in the house for half of one quarter, or about five and one-half weeks. Each student is given primary responsibility for the care of the infant over a four- or five-day period. With a new group of students entering the house at mid-quarter, the infant is subjected to a noncontinuous mothering experience with as many as 16 different mother figures over a period of approximately three months. In addition, there is the discontinuity associated with moving the infant either from the hospital or from a foster home to the home management house, and from the home management house to either a foster home or an adoptive home.

In research currently in progress, the experimental group consists of approximately 40 infants who have lived in or will live in a home management house for a period of one college quarter, beginning in January, 1956. Two other groups of infants, representing two different classes of experimental control, are also being observed. The first of these is a group of infants, matched for age and sex with the experimental group, who are subjected to foster home care for a period of time during early infancy, followed normally by placement in a private adoptive home. The second of these control groups consists of infants born into complete families who remain with their natural parents and experience no unusual mothering situation. Infants in this group are also matched for age and sex with the infants in the experimental group. For convenience, the three groups have been identified as follows: Group A—home management house infants, Group B—infants subjected to foster home care, Group C—infants subjected only to "normal" mothering in the homes of their natural parents.

Infants comprising Groups A and B are brought into the research through the services of the Iowa Children's Home Society. They are children of unmarried mothers. Efforts are made to control any selective factors which might operate to determine which infants are placed in Group A and which in Group B. Thus, it may reasonably be assumed that no important differences exist between the two groups at the time of entry into the research project. Infants in Group C are brought into the research through the cooperation of a pediatric clinic in Des Moines, Iowa.

All infants in the three groups are given an initial testing shortly after birth, one purpose of which is to determine whether they can be accepted as research subjects. Infants with obvious congenital defects such as spastic

paralysis, congenital malformations of limbs or body organs, micro- or hydrocephaly, or serious sensory defects, are excluded. Premature infants are also excluded.

At three months of age each infant is given a complete test battery, which is then administered again at each of the following ages: 6 months, 9 months, 12 months, 18 months, and 24 months. This test battery is given as nearly as possible to the exact day representing each of these ages. However, for practical reasons and difficulties of scheduling, a period of time from two weeks prior to that day to two weeks following that day has been allowed.

The kinds of data being collected on the infants may be classified as follows: (a) medical, (b) psychometric, and (c) anthropometric. The medical data consist of the birth history and reports made regularly by the examining physician. The psychometric data are derived through administration of the Gesell Developmental Schedule, the Cattell Infant Intelligence Scale, and the California Infant Scale of Motor Development. In addition, observational data representing the impressions of the psychological examiners are recorded during the examination period. The development of basic reflex patterns is also noted and recorded. At the time of the 12-months examination a brief motion-picture sequence is made of the infant in a relatively standardized situation. This sequence is taken immediately following the completion of the anthropometric measurements, while the infant is being dressed. This provides a visual record of each subject at 12 months of age, made under comparable conditions for all infants.

Anthropometric measurements recorded at each of the examination ages are: head circumference, chest circumference, calf circumference, biiliac width, thickness of subcutaneous fatty tissue in calf of left leg, over-all length, and weight.

The medical data are obtained by a physician within a few days of the psychological examination. Psychological data are obtained by a team of examiners, usually two persons working together. Anthropometric measurements are taken by a nurse and an assistant, both of whom have received instruction in the techniques of taking physical measurements of infants.

The psychological and anthropometric measurements have been made, in the majority of cases, either in a testing room at the Iowa Children's Home Society in Des Moines, or in a testing room of the Child Development Laboratories at Iowa State College. For the Group C infants the initial examination has been conducted in one of the examining rooms of a pediatric clinic in Des Moines; for the subsequent tests the Group C infants are brought by their mothers to the Iowa Children's Home Society. A few infants in Groups A and B, who are generally adopted before the 6-months test, have sometimes been examined in their private adoptive homes if for some reason the parents could not bring them to one of the testing centers. In such cases the examining teams have attempted to reproduce as many of the testing room conditions as the home situation would

permit. The administration of the complete test battery, excluding the medical examination which is given separately by the physician, requires about one hour, depending on the age of the infant, the ease with which rapport is developed, and other miscellaneous factors.

At the time of this writing, there are 101 subjects in the noncontinuous mothering research project, about evenly distributed among the three groups. A preliminary comparison of 36 infants in Group C with 60 infants in the combined Groups A and B indicates that the home management house and foster home babies are in no way inferior to the natural home babies at the time of induction into the research.

As the data are analyzed, reports will be made which describe more fully the comparative status of Groups A, B, and C at the time of initial examination, as well as the relative progress of each of the three groups during the first two years of life. It is to be hoped that follow-up studies of these same subjects will permit investigation of the long-range effects, if any, of the kind of discontinuous mothering described here.

References

1. Ainsworth, M. D., & Bowlby, J. Research strategy in the study of mother-child separation. *Courrier*, 1954, 4, 105-130.
2. Aubry, J. The case of Monique. In K. Soddy (Ed.), *Mental health and infant development*. Vol. I. New York: Basic Books, 1956. Pp. 125-130.
3. Bakwin, H. Emotional deprivation in infants. *J. Pediat.*, 1949, 35, 512-521.
4. Bender, L. *Aggression, hostility and anxiety in children*. Springfield, Ill.: Charles C. Thomas, 1953.
5. Bowlby, J. *Maternal care and mental health*. Geneva: Monogr. World Health Organization, 1952.
6. Bowlby, J. Mother-child separation. In K. Soddy (Ed.), *Mental health and infant development*. Vol. I. New York: Basic Books, 1956. Pp. 117-122.
7. Bowlby, J., & Robertson, J. A two-year-old goes to hospital. In K. Soddy (Ed.), *Mental health and infant development*. Vol. I. New York: Basic Books, 1956. Pp. 123-124.
8. Dennis, W. Infant development under conditions of restricted practice and of minimum social stimulation: A preliminary report. *J. genet. Psychol.*, 1938, 53, 149-157.
9. Freud, Anna, & Burlingham, Dorothy. *Infants without families*. New York: International Universities Press, 1944.
10. Goldfarb, W. Infant rearing and problem behavior. *Amer. J. Orthopsychiat.*, 1943, 13, 249-265.

11. GOLDFARB, W. Infant rearing as a factor in foster home replacement. *Amer. J. Orthopsychiat.*, 1944, 14, 162-166.

12. GOLDFARB, W. Effects of early institutional care on adolescent personality: Rorschach data. *Amer. J. Orthopsychiat.*, 1944, 14, 441-447.

13. GOLDFARB, W. Effects of psychological deprivation in infancy and subsequent stimulation. *Amer. J. Psychiat.*, 1945, 102, 18-33.

14. GOLDFARB, W. Variations in adolescent adjustment of institutionally reared children. *Amer. J. Orthopsychiat.*, 1947, 17, 449-457.

15. GOLDFARB, W. Emotional and intellectual consequences of psychologic deprivation in infancy: A revaluation. In P. H. Hoch and J. Zubin (Eds.), *Psychopathology of childhood*. New York: Grune & Stratton, 1955. Pp. 105-119.

16. HEINICKE, C. M. Some effects of separating two-year-old children from their parents. *Hum. Relat.*, 1956, 9, 105-176.

17. LOWREY, L. G. Personality distortion and early institutional care. *Amer. J. Orthopsychiat.*, 1940, 10, 576-585.

18. RHEINGOLD, HARRIET L. Mental and social development of infants in relation to the number of other infants in the boarding home. *Amer. J. Orthopsychiat.*, 1943, 13, 41-44.

19. RIPIN, ROWENA. A comparative study of the development of infants in an institution with those in homes of low socio-economic status. *Psychol. Bull.*, 1933, 30, 680-681.

20. SPITZ, R. A. Hospitalism. *Psychoanal. Study Child*, 1945, 1, 53-73.

21. SPITZ, R. A. Anaclitic depression. *Psychoanal. Study Child*, 1946, 2, 313-342.

22. SPITZ, R. A. Hospitalism. *Psychoanal. Study Child*, 1946, 2, 113-117. 113-117.

23. STONE, L. J. A critique of studies of infant isolation. *Child Develpm.*, 1954, 25, 9-19.

A NOTE ON THE SIMILAR EFFECTS OF LOW SOCIAL AVAILABILITY OF AN ADULT AND BRIEF SOCIAL DEPRIVATION ON YOUNG CHILDREN'S BEHAVIOR

JACOB L. GEWIRTZ
Laboratory of Psychology
National Institute of Mental Health

DONALD M. BAER
University of Washington

CHAYA H. ROTH
The University of Chicago

In two recent researches experimental conditions were devised to affect the behavior of young children for social reinforcers from adults. These conditions (which were implemented by different operational procedures in each case) appear to have a number of possibly critical formal characteristics in common. A comparison of them may prove fruitful.

PROCEDURES

In one study (2), children (*Ss*) ranging in age from 4-0 to 5-7 engaged individually in easel painting in the presence of an adult. They could emit attention-seeking responses to the adult under conditions of the apparent low and high availability of the adult as a source of social reinforcers. With one-half (28) of the *Ss*, the adult appeared continually available (High Availability); that is, he sat passively immediately behind the child, attending completely to him. The child was told at the outset that the adult would watch him paint. With the remaining 28 *Ss*, the adult sat at a desk farther away, apparently engrossed in paper work (Low Availability). Here the child was informed at the outset that the adult had work to complete at his desk, but that upon request he would be happy to supply all the materials the child needed. Under both conditions the adult responded briefly to any direct initiation of the child in a friendly and pleasant manner.

In the second study (4), 32 children in a similar age group (3-10 to 5-3) played a two-response marble game individually in the presence of an adult experimenter (*E*). By reinforcing the initially less frequent response with *approval* (i.e., by saying "Good," "Um-hmm," and the like), the relative frequency of the response could be increased. Each child played the game in the adult's presence on two occasions. One, defined as *Social Deprivation*, was preceded by a 20-minute period of social isolation. During

this deprivation period the child waited alone, in a bare room while the game was "repaired" by the adult in an adjoining room. On the other occasion, defined as *Non-deprivation,* the child played the game immediately upon his arrival from his play group.

Factorial control was exercised over the sex of the children and adults in both studies (and different *Ss* and *Es* were employed in each). That is, half of the subjects of each sex interacted individually with a man, the other half with a woman.

Results

Low Availability of the adult and Social Deprivation both effected similar increases in the incidence of behaviors by the children for social reinforcers from the adults. The mean frequency of attention-seeking responses was greater under Low Availability than under High Availability. And, similarly, the mean relative frequency of behaviors for approval was reliably increased by the Deprivation condition, relative to the Non-deprivation condition.

In both studies a similar cross-sex pattern of behavior for a social reinforcer was observed: Children appeared to emit more reinforcer-oriented responses to adults of the opposite sex than to adults of the same sex as themselves. In both studies, this tendency was more marked for boys than for girls. Furthermore, this pattern uniformly qualified the experimental conditions determining the relative supply of these reinforcers to the children, in both studies. That is, there was a suggestion that whether or not *Ss* displayed more attention-seeking to adults of the opposite sex depended on whether the adult was under Low or High Availability. Similarly, the effect of Social Deprivation on what may be called approval-seeking was qualified by a reliable second-order interaction effect between the condition of Deprivation, the Sex of the Child, and the Sex of the Adult.

Figure 1 demonstrates the relative *change* in behavior for social reinforcers, which can be attributed to Low Availability in the one case and to Social Deprivation in the other, as reflected in the four conditions provided by the combinations of the variables Sex of Child and Sex of Adult. That is, increases in behavior for attention from the High Availability condition to the Low Availability condition are compared with increases in behavior for approval from Non-deprivation to Deprivation. The nature of the effect is seen to be remarkably similar in the two cases. If the four Availability mean differences are ranked from high to low, and compared to such rankings of the four Deprivation mean differences, we find that an identical pattern emerges: the rank order is Boys with Women, Girls with Men, Girls with Women, and Boys with Men. The Spearman rank correlation coefficient between these two patterns is $+1.00$ ($p < .05$, single-tail test). From this we may conclude that the concordance between the two patterns is reliable.

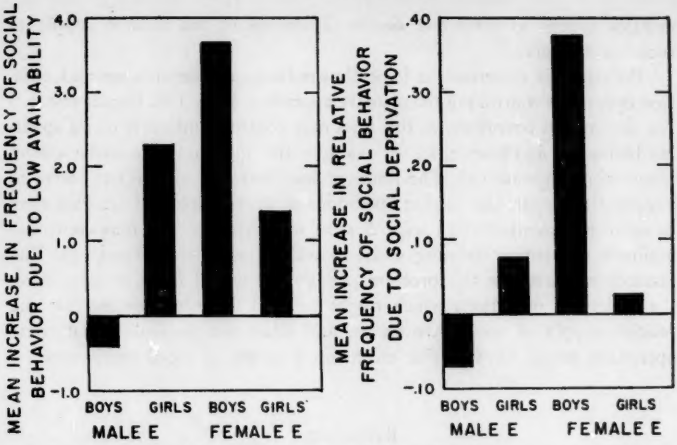

FIGURE 1—A comparison of the mean increases in behaviors for social reinforcers due to Low Adult Availability with those due to Social Deprivation, as reflected against the four Sex of *S* by Sex of *E* conditions.

COMPARISON

This uniform cross-sex qualification of the Availability and Deprivation variables might be referred to a number of rationales [summarized in (2)], the most prominent probably being Freud's Oedipus theory. However, the underlying nature of this interaction is less important here than the fact that it displays such a remarkably similar form under both experimental conditions. This similarity cannot readily be ascribed to chance; consequently, it presents a clear invitation to analyze the formal characteristics of the deprivation and availability conditions. If these conditions can increase behavior for what are presumed to be different social reinforcers in a similar manner under comparable conditions, it is reasonable to suspect that they share important characteristics.

An examination of these two experimental conditions bears out this expectation. Social Deprivation clearly involves a complete limitation of access to all sources of social reinforcement. A fair restatement of this would be that all reinforcing persons are rendered unavailable for a period of time. But the Availability conditions certainly appear to represent two points along precisely this dimension: Low Availability of the adult implies a more limited supply (although not a zero supply) of reinforcers from him than does High Availability.[1] In other words, both experimental pro-

[1] It should be noted that in the Deprivation condition, reinforcers were delivered *after* a period of zero supply of reinforcers; in the Low Availability condition, reinforcers were delivered *during* a period of apparently limited supply of reinforcers.

cedures appear to involve a simple dimension of the relative supply of social reinforcers.

Perhaps this dimension is logically similar to the deprivation and satiation operations controlling the primary appetitive drives (6). Recent research has shown that operations of this class may control reinforcers of no apparent biological significance, as, for example, the opportunity to make a brief observation response (1). The findings described here and in (5)[2] seem to suggest that apparently similar operations of deprivation and satiation may prove to be extendable to a wide class of social stimuli, and may constitute a simple convention defining social as well as appetitive drives (3). The research required for this problem presumably would focus initially upon the variety of operations which might be used to restrict or increase the relative supply of social reinforcers, and upon the possibility that these operations would show similar effects on a variety of social reinforcers.

References

1. Butler, R. A. The effect of deprivation of visual incentives on visual exploration in monkeys. *J. comp. physiol. Psychol.*, 1957, 50, 177-179.
2. Gewirtz, J. L. Three determinants of attention-seeking in young children. *Monog. Soc. Res. Child Develpm.*, 1954, 19, No. 2 (Serial No. 59).
3. Gewirtz, J. L. A program of research on the dimensions and antecedents of emotional dependence. *Child Develpm.*, 1956, 27, 205-221.
4. Gewirtz, J. L., & Baer, D. M. The effect of brief social deprivation on behaviors for a social reinforcer. *J. abnorm. soc. Psychol.*, 1958, 56, 49-56.
5. Gewirtz, J. L., & Baer, D. M. Deprivation and satiation of social reinforcers as drive conditions. *J. abnorm. soc. Psychol.*, in press.
6. Skinner, B. F. *Science and human behavior.* New York: Macmillan, 1953.

[2] A condition of *satiation* for social reinforcers (5) has produced effects consistent with this logic.

A SHORT FORM OF THE CHILDREN'S MANIFEST ANXIETY SCALE

NISSIM LEVY
Brown University

Recently Castaneda *et al.* have published a children's form (CMAS) of the Taylor manifest anxiety scale (1). In the process of carrying out a large scale standardization and item analysis study on subjects from two towns in Rhode Island, we found 10 items which showed no significant differences for either grade or sex.

It was the purpose of the present study to determine the relationship between this 10-item form of the CMAS and the original 42-item form. All protocols were rescored using the 10-item form. This was accomplished very simply since all of the protocols had previously been coded on IBM cards for the item-analytic study.

Since the Rhode Island data indicated that not only were there significant sex and grade differences but also a significant interaction for full-scale scores, separate Pearsonian correlation coefficients were computed for each of the six subgroups which comprised our total population of subjects. Table 1 presents these values.

It can be seen that the short-form scores do quite a respectable job of predicting the full-form scores. The correlations range in value from .84

TABLE 1

SHORT FORM – FULL FORM CORRELATIONS FOR DIFFERENT GRADE LEVELS AND FOR BOYS AND GIRLS SEPARATELY

	GRADE					
	4TH		5TH		6TH	
	N	*r*	*N*	*r*	*N*	*r*
Boys	145	.84	129	.85	151	.88
Girls	130	.84	113	.84	185	.95

to .95. Approximately 70 to 90 per cent of the variance in the full-scale scores can be accounted for by the variance of the scores on the short form.

Another of the possible advantages of the short form lies in the fact that it is composed of items which do not seem to show significant sex or grade differences. One could, it would seem, more reasonably ignore *Ss'* sex or grade in the fourth to sixth grade range when using the short-form scale than when using the full scale.

For those interested in using the short form, the 10 items which comprise it are numbers 2, 9, 13, 15, 18, 26, 31, 43, 44, and 46 as presented in Castaneda *et al.* (1).

References

1. Castaneda, A., McCandless, B. R., & Palermo, D. S. The children's form of the manifest anxiety scale. *Child Develpm.*, 1956, 27, 317-326.

DIFFERENCES BETWEEN MOTHERS' AND FATHERS' RATINGS OF LOW ANXIOUS AND HIGH ANXIOUS CHILDREN[1]

KENNETH S. DAVIDSON, SEYMOUR B. SARASON, FREDERICK F. LIGHTHALL, RICHARD R. WAITE, *and* IRVING SARNOFF
Yale University

As part of an ongoing project concerned with the development and correlates of children's anxiety, this study focuses on differences among parents' ratings of high anxious (HA) and low anxious (LA) children. Recent reports (1, 4, 5) from this project described the Test Anxiety (TA) and General Anxiety (GA) Scales for children[2] which were used to determine placement of children in the LA and HA groups in the present study. The previous reports have presented evidence pertaining to the validity of the anxiety scales. In the present study, validity of the scales was also of primary importance. That is, it was predicted that on items of a personality check list parents of LA children would rate their children more favorably than would parents of HA subjects. In addition, the present procedure permitted another important investigation. Most researchers in this area have relied on mothers' judgments of their children's characteristics, whereas both parents' evaluations of their child were obtained in this study. As a result, it was possible to compare the fathers' ratings with the mothers'. A further refinement was possible in that boys and girls were represented equally in the two anxiety groups. Thus, fathers' and mothers' ratings of LA and HA boys could be investigated as well as both parents' ratings for girls. While no specific differences were expected between fathers' and mothers' judgments, it is clear that any significant differences would have important implications for research relying on parental evaluations of children's personalities.

METHOD

Subjects

The children in this study were selected from 747 pupils in grades two through five who had answered the TA and GA questionnaires in the previous grade. Additionally, all had taken the Otis Alpha Intelligence

[1] The project of which this study is a part is being supported by a grant from the United States Public Health Service (M-712).

[2] Copies of these two scales are available for research purposes only.

Scale. Children were excluded whose parents were separated or divorced or who were regarded as arithmetic, reading, or behavior problems by their teachers. Subjects were considered as HA whose TA and GA scores were in the upper quartile of the distributions, and those whose scores fell in the lower quartiles were placed in the LA group. Thirty-two pairs of children were obtained by matching LA with HA subjects on grade, sex, and Otis Alpha T scores. Four such pairs were matched for each grade-sex group and the design was a 2 (anxiety) by 2 (sex) by 4 (grade) arrangement, with four subjects in each cell.

Procedure

The parents' ratings of their child were obtained by one of us (K.S.D.), who interviewed them in their homes at their convenience. In the first interview, the mother responded to questions about the child's school experiences, and the parents filled out an attitude scale on child rearing. A second interview with the mother followed about two months later and pertained to the child's experiences as well as personality in general. After the first item, a general free description of the child, the parent or parents filled out the personality check list.

This check list consisted of 25 items representing a variety of personality characteristics. For each item, a rating was required on a six-point scale with definitions for the extremes provided. For example, item 1 was presented as follows:

Immature	3	2	1	1	2	3	Mature
Acts younger than age							Acts grown up

Each parent was instructed to compare their child with children of his own age. They were to decide which of the two words described him better. Then they circled one of the three numbers on that side, with the three numbers defined as 1, a little more on this side; 2, definitely on this side; and 3, very much on this side. There were 16 items with clearly favorable and unfavorable extremes which were placed randomly in the list with the favorable extremes put on the right and left in random fashion.

If the father was not present, the check list was left for him with appropriate instructions and the explanation that the mother had expressed her own opinions and that she should not influence the father's rating of the child. In most cases the fathers were not present, but cooperation was excellent and the ratings by both parents of the 64 subjects were obtained.

Results

All analyses of the evaluations were directional only and did not take account of the size of the difference between ratings. The procedure involved comparing the ratings of each matched pair of subjects on each of

the 16 items previously judged by three people as clearly having a favorable value at one extreme and an unfavorable value at the other. For the 16 pairs of HA and LA boys, the fathers' ratings of each matched pair were examined to see whether the father of LA rated his son more favorably or unfavorably than the father of the HA judged his boy. For each item, the number of times the father of the LA boy rated him more favorably than the HA boy's father judged his son (matched with the LA son) was counted as well as the number of times the LA boy was judged less favorably than his mate. The obtained proportion was subjected to a sign test and the p value was computed. This test was used for each of the 16 items. The instances in which the fathers of a matched pair of LA and HA boys rated their sons equally favorably were not considered in obtaining the

TABLE 1

DIFFERENCE IN PARENTS' FAVORABLE RATINGS OF LA AND HA CHILDREN*

Item†	*Father –Boy*	*Mother –Boy*	*Father –Girl*	*Mother –Girl*	*Parent –Boy*	*Parent –Girl*	*Father –Child*	*Mother –Child*	*Parent –Child*	*Mean Difference Test*‡
Mature	.02	.29	.29	.78	.02	.50	.02	.50	.07	*.025*
Sociable ...	.29	.50	.50	.50	.35	.50	.35	.50	.44	*.05*
Unanxious .	.28	.50	.30	.30	.42	.28	.16	.43	.18	*.001*
Optimistic ..	.39	.39	.07	.50	.27	.15	*.08*	.42	.12	*.001*
Responsible	.07	.50	.09	.50	.15	*.10*	.02	.50	*.04*	*.005*
Active	.71	.50	.50	.75	.75	.76	.75	.76	.87	*.995*
Relaxed ...	.60	.71	.13	.29	.77	.09	.36	.50	.39	.20
Leads Others	.19	*.06*	.50	.50	*.03*	.50	.27	.20	.12	*.005*
Independent	.61	.50	.71	.50	.65	.58	.77	.50	.71	*.995*
Generous ..	.50	.81	.97	.72	.74	*.98*	.92	.90	*.98*	*.999*
Ambitious .	.11	.13	.50	.78	*.03*	.60	.13	.41	.14	*.05*
Affectionate	.87	.77	.83	.39	.92	.59	.95	.50	.88	*.995*
Adaptable .	.81	.60	.13	.19	.83	*.06*	.50	.43	.38	.25
Care in Plans	.28	.13	.39	.81	.14	.58	.29	.50	.27	*.10*
Not Sensitive	.40	.29	.15	.50	.22	.16	.28	.27	.09	*.001*
Takes Blame	.39	.61	.50	.50	.50	.42	.34	.50	.44	.15
Mean Difference Test§	*.01*	.25	.15	.60	.20	.20	*.10*	.50	*.10*	*.002*¶

* Probabilities below .50 indicate that the LA were rated more favorably than the HA and values over .50 indicate that the HA were judged more favorably than the LA subjects. Items significant in the predicted direction at or below $p = .10$ are italicized as well as those items significant in the opposite direction with a p value of .90 or above.

† The terms for the favorable extremes of the items are listed.

‡ $N = 9$.

§ $N = 16$.

¶ $N = 144$.

probabilities, since each pair of LA and HA children was matched carefully on intelligence, sex, and grade. Following the same procedure, the ratings of the LA subjects were compared with those of the HA subjects for the various parent-child groupings given in the first row of Table 1. In all the groupings, mothers' ratings were compared with mothers', and fathers' compared with fathers'.

The results of these analyses are presented in Table 1 in the form of probabilities. For each parent-child grouping, there is a column of p values, one for each of the 16 items. At the bottom of each column, there is a significance point determined by a t test computed to evaluate the mean difference from chance of the 16 p values in each column. In addition, for each item, t tests were computed to determine the significance of the mean difference from chance of the nine probabilities obtained for the various breakdowns of parent-child groups (column 10). Finally, the mean difference test was done for all 144 p values in the main body of the table.

For all 144 p values, the mean difference from .50 is significant at almost the .002 level. The same results are reflected in the column of parents' ratings of children, where only four items yield results in the wrong direction while 12 are in the predicted direction, and the mean difference test is significant below the 10 per cent point of confidence. Of the nine mean difference tests for the columns, seven are in the predicted direction. Over-all, the evidence indicates that parents of LA children rate them more favorably than parents of HA children rate their children on these variables.

The results of the comparison of ratings of LA by their mothers with ratings of HA by their mothers were pooled with the results of the comparison of fathers' ratings of LA with ratings of the HA by their fathers. When this is done, the results tend to significance in the predicted direction (column 9). When the judgments of parents are pooled, their ratings of LA boys and of LA girls are more favorable, though not significantly, than parents' judgments of HA boys and HA girls (columns 5 and 6). However, in comparing fathers' with mothers' evaluations, significance is attained and approached only for the fathers. According to the mean difference test, the fathers of LA subjects rate them more favorably than the fathers of HA subjects rate them. That result is significant for the fathers-boys column and tends to significance for the fathers-children and fathers-girls categories. By that test, the mothers of LA do not judge them more favorably than mothers of HA evaluate their children beyond chance, and in fact, to a slight degree, mothers of HA girls rate them more favorably than their LA mates are judged by their mothers. In other words, HA and LA children tend to be discriminated in the predicted direction by fathers but not by mothers.

When attention is focused on individual items, there are three items on which fathers discriminate between HA and LA in the predicted direction, while mothers do not (columns 7 and 8). There is another item on which fathers also discriminate between HA and LA but not in the pre-

dicted direction. In short, whatever discriminations are made are the results of the fathers' ratings and not the mothers'.

In column 5 are given the results of a comparison of the ratings by parents of HA and LA boys. The corresponding results for girls are in column 6. When these two columns are viewed in relation to each other, two things seem worthy of note. First, there are certain items (mature, leads others, generous, affectionate) in which HA and LA boys are discriminated but girls are not, or HA and LA girls are discriminated but boys are not. Second, there are two items (relaxed and adaptable) in which LA girls are rated more favorably than HA girls, while on the same items HA boys are rated more favorably, though not significantly, than LA boys. There is also one item (ambitious) on which LA boys are rated more favorably than HA boys, whereas HA girls are rated slightly more favorably on that item than LA girls. Another way of stating the above is that there is not a single item which discriminates significantly between HA and LA boys as well as between HA and LA girls. This finding does not permit of any clear-cut conclusion. It does suggest, however, that the anxious boy as defined in this project is perceived by parents differently than the anxious girl.

Discussion

The most suggestive finding reported above concerns the failure of mothers' ratings to discriminate between the anxiety groups whereas the fathers' ratings do discriminate in the predicted direction. This result takes on added significance in light of the fact that significant and predicted differences between these 32 matched pairs of children have been obtained in other situations: direct classroom observations, the Rorschach situation, and an experimental learning task[3] (3, 4, 5).

Several explanations, of course, are possible. It may be that fathers, who spend less time with their children than do mothers, are less emotionally involved with their children and perceive them somewhat more dispassionately. It may be also that the mothers, who previous to the ratings had been asked many questions about the adjustment of their child, were in some way compensating for or reacting against whatever facts or attitudes they had revealed to the interviewer. It is also conceivable that the ratings of the mothers and fathers are not discrepant in that they may be using samples of behavior which differ in terms of time (e.g., day vs. night), content (e.g., observation of peer relationships) and variety of possible reactions. It is one thing to say that the ratings of mothers and fathers are different and another thing to say that one is valid and the other is not. The only conclusion we feel entitled to draw is that developmental studies which

[3] In the study involving the experimental learning task, 24 of the 32 matched pairs of children were used for reasons cited in that publication.

rely on the ratings or observations by the mother, and this characterizes almost all such studies, may in overlooking the father be robbing themselves of important data.

It was noted earlier that while there were certain items which discriminated between HA and LA boys and others which discriminated between HA and LA girls, there was not a single item which discriminated for both sexes. We interpret this as suggesting that anxiety, as defined by scores on our two scales, has different psychological significance for boys and girls. The following very brief summary of findings of other of our studies may clarify this interpretation. Although girls consistently get higher scores than boys on both the Test Anxiety and General Anxiety Scales, the predicted interfering effects of anxiety tend to show up much more clearly between HA and LA boys than between HA and LA girls. We have elsewhere (4) suggested that our culture makes it easier for a girl to admit to anxiety than for a boy—one might say that we expect and support the admission of anxiety in a girl far more than in a boy. We tentatively hold to the opinion that boys who score high anxious are admitting to reactions which they experience as ego-alien to a greater degree than is true of girls. This opinion is not incompatible with the finding in the present study that no item of the rating scales discriminated both between HA and LA boys and between HA and LA girls.

References

1. Sarason, S. B., Davidson, K. S., Lighthall, F. F., & Waite, R. R. A test anxiety scale for children. *Child Develpm.*, 1958, 29, 105-113.
2. Sarason, S. B., Davidson, K. S., Lighthall, F. F., & Waite, R. R. Direct classroom observations of high anxious and low anxious children. Unpublished data.
3. Sarason, S. B., Davidson, K. S., Lighthall, F. F., & Waite, R. R. Rorschach behavior and performance of high anxious and low anxious children. *Child Develpm.*, in press.
4. Sarnoff, I., Lighthall, F. F., Waite, R. R., Davidson, K. S., & Sarason, S. B. A cross-cultural study of anxiety amongst American and English school children. *J. educ. Psychol.*, in press.
5. Waite, R. R., Sarason, S. B., Lighthall, F. F., & Davidson, K. S. A study of anxiety and learning in children. *J. abnorm. soc. Psychol.*, in press.